

STAT

Page Denied

UNCLASSIFIED

STAT

AEC-tr-2876 (Pt. I)

UNITED STATES ATOMIC ENERGY COMMISSION

QUANTUM ELECTRODYNAMICS

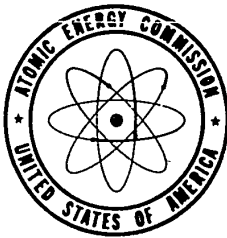
(Part I)

A. I. Akhiezer and V. B. Berestetsky

Translated by:

CONSULTANTS BUREAU, INC.

Technical Information Service Extension, Oak Ridge, Tenn.



UNCLASSIFIED

A translation of:

Kvantovaya Elektrodinamika, A. I. Akhiezer and V. B. Berestetsky Moskva,
Gosudarstvennoe Izdatel'stvo Tekhniko-Teoreticheskoi Literatury, 1953, 428 pp.

Printed in the U.S.A. Price \$ 2.65. Bound in two parts to be sold together.
Available from the Office of Technical Services, Department of Commerce,
Washington 25, D.C.

$$\left. \begin{aligned} (\omega - m) \varphi(\mathbf{k}) - \sigma \mathbf{k} \chi(\mathbf{k}) &= 0, \\ -\sigma \mathbf{k} \varphi(\mathbf{k}) + (\omega + m) \chi(\mathbf{k}) &= 0. \end{aligned} \right\} \quad (8.6)$$

The set of equations (8.6) can be solved if

$$\begin{vmatrix} \omega - m & -\sigma \mathbf{k} \\ -\sigma \mathbf{k} & \omega + m \end{vmatrix} = \omega^2 - m^2 - k^2 = 0, \quad (8.7)$$

so that the frequency ω is given by

$$\omega = \pm \epsilon, \quad (8.8)$$

where

$$\epsilon = \sqrt{k^2 + m^2}. \quad (8.9)$$

Equation (8.9) expresses the well-known relativistic relation between the energy and momentum.

Thus, there exist two kinds of solutions for the Dirac equation corresponding to the two signs in Equation (8.8). We shall call these positive- and negative-frequency solutions. The general solution to the Dirac equation can then be written

$$\psi = \psi^{(+)} + \psi^{(-)}, \quad (8.10)$$

where

$$\left. \begin{aligned} \psi^{(+)} &= \int \psi_0^{(+)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r} - i\epsilon t} d\mathbf{k}, \\ \psi^{(-)} &= \int \psi_0^{(-)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r} + i\epsilon t} d\mathbf{k}. \end{aligned} \right\} \quad (8.11)$$

A solution whose frequency has a definite sign shall also be written $\psi^{(\eta)}$, where $\eta = \frac{\omega}{\epsilon} = \pm 1$.

Of the two spinors that go to make up $\psi^{(\eta)}(\mathbf{k})$, namely $\varphi^{(\eta)}(\mathbf{k})$ and $\chi^{(\eta)}(\mathbf{k})$, one is an arbitrary

A. I. Akhiezer and V. B. Berestetsky

QUANTUM ELECTRODYNAMICS

State Technico-Theoretical
Literature Press
Moscow, 1953

NOTE

This monograph is an orderly development of quantum electrodynamics based on the most recent works in this field. It includes both the elements of the theory and numerous applications to the calculation of various effects.

The book is intended for readers with sufficient mathematical background who are familiar with quantum mechanics.

The book will be useful for theoretical physicists and for students of modern theoretical physics.

PREFACE

The concepts of particle and field are fundamental to modern physics. In an earlier phase of physics, the so-called classical, these two concepts referred to different physical objects; for instance, the first referred to the electron, and the second to light.

The further development of physics, however, showed that these concepts reflect different aspects of one and the same object. The electron has wave properties, and light (the electromagnetic field) manifests those of a particle, namely the photon. Just as electrons interact by means of the electromagnetic field, so do photons interact by means of the electron-positron field.

Quantum theory shows the unity of the corpuscular and wave aspects of physical objects. The concepts of particle and field join in the unified concept of the quantum field which, for example, makes it possible to describe the processes of production and annihilation of particles within the framework of the existing theory.

At the present time many particles are known, and to these correspond various interacting quantum fields. Of the many forms of physical interaction existing in nature, however, at the present time only the gravitational and electromagnetic interactions have been studied in sufficient detail. The theory of the latter interaction is the subject of quantum electrodynamics, to the systematic exposition of which this book is devoted.

Since electromagnetic interactions are fundamental for the electron and photon, quantum electrodynamics makes it possible to explain and predict a wide range of phenomena related to the behavior of these particles. As for the application of quantum electrodynamics to other particles (nucleons and mesons), it is extremely restricted due to the essential role played by other types of interactions (nuclear or meson interactions) for those particles. Therefore, meson problems are not treated in this book, and the interaction of nucleons with the electromagnetic field is treated only in the low-velocity limit.

The formulation of the fundamental equations of quantum electrodynamics, and even the possibility of separating the interacting fields into a weak one. This situation is expressed by the small magnitude of the constant $\alpha = e^2 / \hbar c$ which characterizes the interaction. Thus, the interaction between the fields in quantum electrodynamics is treated as a small perturbation, and the mathematical method used is perturbation theory in which all quantitative results are presented in terms of power series in α .

Since the electromagnetic and electron-positron fields are systems with unlimited degrees of freedom, the application of perturbation theory gives rise to divergent expressions characteristic of the modern theory, which are absent only in the first nonvanishing approximation of perturbation theory. The development of quantum electrodynamics in recent years made it possible to establish principles for regularizing divergent expressions, so that it then became possible to calculate higher approximations (the so-called radiative corrections).

This progress is to a great extent due to the new, invariant formulation of perturbation theory. Invariant perturbation theory made it possible to represent the results in a compact and relativistically invariant form, which allowed a formulation of the rules for removing singularities. Furthermore, invariant perturbation theory has significant practical advantages over the earlier methods even for first-order calculations. Therefore, the whole exposition in this book is constructed on the basis of invariant perturbation theory.

Although it is a completely satisfactory theory in a definite field of physical phenomena, modern quantum electrodynamics has important drawbacks in that it necessitates the introduction of additional concepts which are neither contained in the fundamental formulation of the theory nor reflected in its basic equations; these concepts are necessary in order to remove the divergences which arise in the theory. This state of affairs seems

to arise from profound causes. They lie in the fact that it is sometimes impossible to construct a closed theory of a limited set of phenomena (in the present case, the pure electromagnetic ones) without accounting also for a wider class of interactions in nature.

The structure of the present book is the following. The first three chapters are devoted to the theory of free, i.e., noninteracting particles (electrons and photons). This corresponds to the above-mentioned fundamental trait of quantum electrodynamics, which makes it possible to consider the interaction as a perturbation. In Chapter III the vacuum is defined as the state of the field in which no particles exist. Nevertheless, the vacuum (as opposed to the metaphysical "void") has physical properties, and it is necessary to take account not only of interaction of particles with each other when considering various phenomena, but also their interaction with the vacuum (this interaction is considered in Chapter VIII). The interacting field equations are formulated in Chapter IV, where invariant perturbation theory is developed. Concrete problems reduce to the calculation of S matrix elements; Chapter V is devoted to a general investigation of this matrix. The rest of the exposition is based on the use of the S matrix. Part of Chapter V (Sections 25-27) is devoted to an analysis of the divergences in the S matrix and to a description of the methods for removing them. The results of these sections are not used in Chapters VI and VII. Therefore, the reader may omit these sections in the first reading, returning to them before going on to Chapter VIII. Chapters VI and VII consider various concrete phenomena in the first nonvanishing approximation. The theory of radiative corrections is developed in Chapter VIII and is based on the methods for removing divergences described in Chapter V. Appendix I describes the general theory of free fields, important special cases of which are the electron-positron field and the electromagnetic field. Appendix II is devoted to the general theory of bound states.

In writing Chapters V and VIII, as well as Sections 32 and 56, we received aid from R. V. Polovin, who performed many of the calculations. G. Ya. Lyubarsky and L. E. Pargamanik participated in writing Sections 48, 50, and 52 (the latter, Section 52). Appendix II was written by A. D. Galanin. We are sincerely grateful to all of them.

We express our gratitude to Academician L. D. Landau, Professor I. Ya. Pomeranchuk, and to the members of the seminars they directed for discussing many problems described in this book.

TABLE OF CONTENTS

PREFACE	3
---------------	---

CHAPTER I

QUANTUM MECHANICS OF THE PHOTON

1. The Photon Wave Function in Momentum Space	11
1. Introduction. 2. Wave Function in $\mathbf{k}$ -space. 3. Energy.	
2. Momentum Eigenstates	16
1. Momentum. 2. On the Photon Wave Function in Configuration Space. 3. Plane Waves.	
3. Angular Momentum. Spin of the Photon	20
1. Angular Momentum Operator. 2. Spin Operator. 3. Spin Wave Functions.	
4. Angular Momentum and Parity Eigenstates	28
1. Angular Momentum Eigenfunction. 2. Spherical Vectors. Parity. 3. Expansion in Spherical Waves. 4. Expressions for the Electric and Magnetic Fields.	
5. Potentials	39
1. Transverse, Longitudinal, and Scalar Potentials. 2. Longitudinal and Scalar Components of the Photon Wave Function. 3. Plane and Spherical Wave Potentials.	
6. The Two-Photon System	45
1. Two-Photon Wave Function. 2. Even and Odd States. 3. Classification of States with a Given Angular Momentum.	

CHAPTER II

RELATIVISTIC QUANTUM MECHANICS OF THE ELECTRON

7. The Dirac Equation	53
1. Spinors and Pauli Matrices. 2. Dirac Equation. 3. On the Necessity for Four-Component Wave Functions. 4. Invariance of the Dirac Equation. 5. The γ Matrices. Continuity Equation. 6. The Transformation Characteristics of Bilinear Combinations of the Wave Functions.	
8. Electron and Positron States	65
1. Solutions with Positive and Negative Frequencies. 2. The Wave Function of the Positron. 3. Positron Parity. 4. Charge-Conjugate Function.	
9. Momentum and Polarization Eigenstates	74
1. Plane Waves. 2. Polarization States. 3. Sum Over Polarizations. 4. Calculation of Traces.	

10. Angular Momentum and Parity Eigenstates of the Electron	85
1. Orbital and Spin Functions. 2. Spherical Spinors. 3. Angular Momentum Eigenfunction. 4. Parity of a State. 5. Expansion in Spherical Waves.	
11. The Electron in an External Field.....	98
1. Dirac Equation in an External Field. 2. Separation of Variables in a Central Field. 3. Asymptotic behavior. 4. Level Behavior as a Function of the Potential Well Depth.	
12. Electron Motion in the Field of the Nucleus.....	105
1. Solution of the Radial Equations for the Coulomb Field. Discrete Spectrum. 2. Wave Functions of the Continuous Spectrum. 3. Isotope Level Shift. 4. General Investigation of the Effect of Finite Nuclear Dimensions. 5. On the Existence of Bound States for Large Z.	
13. Scattering of Electrons	118
1. Born Approximation for Scattering in a Coulomb Field. 2. Scattering in a Central Field. Spinor Scattered Amplitude. 3. Scattering Cross Section in Terms of the Phase Shifts. 4. Azimuthal Asymmetry. 5. Polarization in Scattering. 6. Scattering in a Coulomb Field.	
14. The Nonrelativistic Limit.....	130
1. Transition to the Pauli Equation. 2. Second Approximation. 3. On the Applicability of the Dirac Equation to Nucleons.	

CHAPTER III

QUANTIZATION OF THE ELECTROMAGNETIC AND ELECTRON-POSITRON FIELDS

15. Quantization of the Electromagnetic Field.....	137
1. Four-Dimensional Form of the Field Equations. Plane Waves. 2. Quantum Conditions. 3. Definition of the Vacuum for the Electromagnetic Field. The Use of the Indefinite Metric. 4. Wave Function of a System of Photons in Momentum Space and Second Quantization. 5. Method of Fock Functionals.	
16. Commutators of the Electromagnetic Field. The Singular Functions D , $D^{(1)}$, D^F	162
1. Quantum Conditions for the Potential. 2. Expectation Value of the Operator $\{A_\mu(x), A_\nu(x')\}$ in the Vacuum State. 3. Chronological Product of the Operators $A_\mu(x)$ and $A_\nu(x')$. The D^F function.	
17. Quantization of the Electron-Positron Field.....	173
1. Variational Principle for the Dirac Equation. Energy-Momentum Tensor. 2. Quantum Conditions for the Electron-Positron Field.	

18. Anticommutators of the Electron-Positron Field. The Singular Functions $\Delta(x)$, $\Delta^{(1)}(x)$, $\Delta^F(x)$, $\Delta^R(x)$	181
1. Quantum Conditions for the Operators ψ , $\bar{\psi}$. 2. Definition of the Current. Charge Conjugate Operators. 3. Chronological Products of the Field Operators. 4. Ordered Products of the Field Operators. 5. Representations of the Singular Functions.	

CHAPTER IV

FUNDAMENTAL EQUATIONS OF QUANTUM ELECTRODYNAMICS

19. Interaction between the Electron-Positron and the Electromagnetic Fields	200
1. Fundamental Equations of the Interacting Fields. 2. Variational Principle. 3. Charge Parity. 4. Schrodinger Equation for a System of Fields. "External" Field. "Given" current.	
20. Perturbation Theory. Transition from the Schrodinger Representation to the Interaction Representation	210
1. Interaction Representation. 2. Perturbation Theory.	
21. Covariant Perturbation Theory. Transition from the Heisenberg Representation to the Interaction Representation	214
1. Expansion of the Interacting Field Operators in a Power Series in the Electron Charge. 2. Transition to the Interaction Representation.	

CHAPTER V

THE S MATRIX

22. Calculation of the S-Matrix Elements	230
1. The S Matrix. 2. Matrix Elements of the Field Operators. 3. Representation of the S Matrix as a Sum of Normal Products.	
23. Graphic Representation of the Matrix Elements	240
1. Graphic Representation of Normal Products. 2. Various Field Interaction Processes.	
24. The S Matrix in the Momentum Representation	245
1. General Formula. 2. Example. Furry's Theorem. 3. Summary of the Rules.	
25. Analysis of the Singularities of the S Matrix.....	252
1. General Properties of the Diagrams. 2. Equivalent Skeleton Diagrams. 3. Possible Types of Divergences Related to Irreducible Diagrams.	
26. Removal of the Divergences from the S Matrix.....	260
1. Removal of the Divergence Due to Photon-Photon Scattering. 2. Electron and Photon Self-Energies. 3. Removal of the Divergence Due to the Electron Self-Energy. 4. Removal of the Divergence Due to the Photon Self-Energy. 5. Removal of the Divergence Due to the Vertex Part. 6. Removal of Reducible and Overlapping Divergences.	
27. Mass and Charge Renormalization.....	273
1. The Renormalization Concept. 2. The Relation Between the Divergences Due to the	

Vertex Part and the Electron Self-Energy Part. 3. Removal of Divergences by Means of Auxiliary Masses. 4. On the Divergence of the Renormalized Series for the S Matrix.

28. The Probability for Various Processes..... 285
1. General Probability Formula. 2. Summation and Averaging Over Photon and Electron Polarizations. 3. Probabilities in the Presence of an External Field. 4. Effective Perturbation Energy.

CHAPTER VI

THE INTERACTION OF ELECTRONS AND PHOTONS

29. Scattering of a Photon by an Electron 295
1. Impossibility of First-Order Processes for the Free Electron. 2. Photon Scattering by a Free Electron. 3. Angular Distribution and Total Cross Section for Unpolarized Photons. 4. Angular Distribution for Scattering of Polarized Photons.
30. Emission and Absorption of a Photon 303
1. General Expression for the Matrix Element. 2. Electric Multipole Radiation. 3. Magnetic Multipole Radiation. 4. Selection Rules. Order of Magnitude Evaluations. 5. Absorption of a Photon. 6. Photoelectric Effect.
31. Bremsstrahlung 315
1. General Expression for the Matrix Element. 2. Perturbation Theory for the Wave Function of an Electron in the Continuous Spectrum. 3. Bremsstrahlung Cross Section. 4. Angular Distribution for Radiation in a Coulomb Field. 5. Bremsstrahlung Spectrum. 6. Screening. 7. Radiation From Electron-Electron and Electron-Positron Collisions.
32. Emission of Long-Wavelength Photons 329
1. Infrared "Catastrophe". 2. Investigation of the Divergence in the Low-Frequency Region. 3. The Bloch-Nordsieck Method. 4. Relation Between the Photon "Mass" and the Minimum Frequency.
33. Disintegration of an Electron-Positron Pair Into Photons. Pair Production by Photons 346
1. Two-Photon Annihilation. 2. Positronium Decay. 3. Three-Photon Orthopositronium Decay. 4. Single-Photon Pair Annihilation. 5. Pair Production by a Photon in an External Field. 6. Pair Production in Photon-Electron Collisions.
34. Photon Scattering by a Bound Electron. Emission of Two Photons 356
1. Dispersion Equation. 2. Emission of Two Photons. Metastable $2S_{1/2}$ State in Hydrogen. 3. Resonance Scattering. 4. Angular Correlation in Successive Emission of Two Photons.

CHAPTER VII

RETARDED INTERACTION BETWEEN TWO CHARGES

35. Interaction Function of Two Charges. Retarded Potentials 371
1. Interaction Matrix of Two Charges. 2. General Form of the Matrix Element. 3. Retarded Potentials and Transition Currents.

36. Electron and Positron Scattering by an Electron 380
1. Electron-Electron Scattering. 2. Positron-Electron Scattering.
37. Interaction Energy for Two Electrons Up to Terms in $\frac{v^2}{c^2}$ 385
1. The Breit Formula. 2. The Schroedinger Equation for a Two-Electron System. 3. Electron-Positron Interaction. 4. Exchange Interaction Between an Electron and Positron.
38. Positronium 395
1. The Hamiltonian and Unperturbed Equation. 2. Perturbation Operator. 3. Fine Structure. 4. Zeeman Effect.
39. Internal Conversion of Gamma Rays 407
1. Expansion of Retarded Potentials in Spherical Waves. 2. Conversion Coefficient. 3. K-Shell; Reduction to Radial Integrals. 4. K-Shell; Results. 5. Effect of the Finite Size of the Nucleus.
40. Conversion With Pair Creation. Nuclear Excitation by Electrons 426
1. Conversion of Magnetic Multipole Radiation. 2. Conversion of Electric Multipole Radiation. 3. Nuclear Excitation by Electrons. 4. Monochromatic Positrons. 5. Pair Production in Particle Collisions.
41. 0-0 Transitions 440
1. Reduction to the Static Interaction. 2. Conversion and Nuclear Excitation in 0-0 Transitions.

CHAPTER VIII

RADIATIVE CORRECTIONS. VACUUM POLARIZATION.

42. Third-Order S Matrix 446
1. Third-Order Matrix Elements. 2. Calculation of the Matrix Element for Radiative Corrections to Electron Scattering.
43. Vacuum Polarization 454
1. Calculation of the Vacuum-Polarization Matrix Element. 2. Renormalization of the Matrix Element.
44. Effective Electron Potential Energy. Magnetic Moment of the Electron 460
1. Effective Electron Potential Energy. 2. Radiative Corrections to the Electron Magnetic Moment.
45. Radiative Corrections to Scattering 463
1. Radiative Corrections to Scattering of an Electron by an External Field. 2. Radiative Corrections to Photon-Electron Scattering. 3. Natural Line Width.
46. Radiative Level Shift for Atomic Electrons 475
1. The Interaction of an Atomic Electron with the Zero-Point Field Oscillations. 2. Radiative Atomic Level Shift. 3. Radiative Level Shift in Muonium.

47. Nonlinear Effects in Electrodynamics..... 484
 1. Scattering of Light by Light. 2. Coherent Nuclear γ -ray Scattering. 3. Lagrangian, Including Nonlinear Effects. 4. Concluding Remarks.

APPENDIX

I. THE THEORY OF WAVE FIELDS.

48. The Wave Functions of the Field and Their Transformation..... 495
 1. Wave Functions of the Field and the Lorentz Group. 2. Irreducible Finite-Dimensional Representations of the Lorentz Group. 3. Direct Product of Representations. 4. The Three-Dimensional Rotation Group. 5. Irreducible Finite-Dimensional Representations of the Orthochronous Lorentz Group.
49. The Energy-Momentum Tensor and the Current Vector..... 503
 1. Energy-Momentum Tensor. 2. Current Vector.
50. Relativistically Covariant Field Equations..... 507
 1. General Form of Relativistically Covariant Field Equations. 2. Invariant Lagrangian.
51. The Mass and Spin of a Particle..... 514
 1. Mass and Spin Values for a Given Equation. 2. The Spin Values of Particles Described by a Given Equation.
52. Examples of Wave Equations for Particles of Various Spins..... 515
 1. Wave Equations with Positive-Definite Charge Density. 2. Wave Equations with Positive-Definite Energy Density.
53. Field Quantization: Spin and Statistics..... 521
 1. The Impossibility of Positive-Definite Charge Density for Particles with Integral Spin. 2. The Impossibility of Positive-Definite Energy Density for Particles with Half-Integral Spin. 3. Field Quantization for Integral and Half-Integral Spins.
- II. BOUND STATE EQUATIONS.
54. The Equation of Motion of an Electron in an External Field with Radiative Corrections Taken Into Account..... 527
 1. The Method of Successive Approximations. 2. Electromagnetic Vacuum Expectation Value. 3. Vacuum Polarization. Electron-Positron Vacuum Expectation Value.
55. The Equation of Motion of Two Interacting Electrons with Radiative Corrections Taken Into Account..... 536
 1. The Equation of Motion of an Electron in a Real Photon Field. 2. The Equation of Motion of Two Interacting Electrons.
- III. MATHEMATICAL APPENDIX
56. Calculation of Certain Integrals..... 539
 1. The Calculation of Integrals over a Finite Invariant Region. 2. Summary of the Integrals.
57. L-Vectors and Spherical Functions..... 545
 1. Irreducible Tensors. 2. L-Vector Algebra. 3. Generalized Spherical Functions.

CHAPTER I

QUANTUM MECHANICS OF THE PHOTON

§ 1. The Photon Wave Function in Momentum Space.

1. Introduction

The corpuscular properties of light were historically the first fundamental fact which established the basis for the development of quantum theory. The relation between the energy of the light particle, the photon, and the frequency of the electromagnetic field corresponding to it,¹⁾

$$w = \hbar\omega \quad (1.1)$$

is historically the first relation containing the quantum constant $\hbar$. However, the systematic quantum mechanics of the atom was developed before that of the photon. This situation has deep physical meaning. Atomic particles, the electrons and nuclei, have rest masses different from zero. These particles can possess energies small with respect to their rest energy, and in this energy region relativistic effects may be neglected. At the present time, however, only nonrelativistic quantum mechanics may be considered a relatively complete part of quantum theory. Since the rest mass of the photon vanishes, no nonrelativistic energy region exists for it; the quantum mechanics of the photon must necessarily be relativistic from the very start.

Quantum mechanics replaces the particle by the wave function field ψ , which determines the probability distribution and the expectation values of the various physical quantities referring to the particle. The particle motion is determined by the field equation (Schrodinger equation). The principle of relativity imposes the requirement of Lorentz invariance on this equation. This requirement is not sufficient for a unique choice of the equation which would describe the individual properties of a given type of particle (see Section 50). In the case of the photon, however, the choice is made simpler by the existence of a classical analog (the classical electromagnetic field). It is natural to choose Maxwell's equations as the quantum mechanical equations of motion for the photon; then the wave properties of the photon will be identical with those of the electromagnetic field. We shall see that together with the quantum postulate (1.1) this is sufficient to construct a theory of photons and their interaction with charged particles.

Our first problem is the study of photons in the absence of electric charges. Although it is just in interactions with other particles that the particle properties appear, such considerations are useful as a preparatory phase for the study of interactions, particularly, as it shall turn out, since the latter can be treated by perturbation methods.

We shall henceforth use a system of units in which Planck's constant divided by 2π and the velocity of light are set equal to unity:

$$\hbar = c = 1.$$

In this system of units, Equation (1.1) can be written

$$w = \hbar k, \quad (1.2)$$

where k is the wave number.

¹⁾ [Underlined letters will denote italicized letters in the original Russian — editor's note].

2. Wave Function in k -space.

The electromagnetic field in empty space is described by the vectors $\underline{E}$ (electric field) and $\underline{H}$ (magnetic field) which satisfy Maxwell's equations:

$$\left. \begin{aligned} \text{curl } \underline{E} &= -\frac{\partial \underline{H}}{\partial t}, \\ \text{div } \underline{H} &= 0, \\ \text{curl } \underline{H} &= \frac{\partial \underline{E}}{\partial t}, \\ \text{div } \underline{E} &= 0. \end{aligned} \right\} \quad (1.3)$$

In order to give a "corpuscular" interpretation to Equations (1.3) let us compare this system to Schrodinger's equation of ordinary quantum mechanics; this is conveniently done by first subjecting Equation (1.3) to a Fourier transformation (transition to the space of wave vectors). Let us represent $\underline{E}$ and $\underline{H}$ in the form

$$\left. \begin{aligned} \underline{E} &= \int \underline{E}(\underline{k}) e^{i\underline{k} \cdot \underline{r}} d\underline{k}, \\ \underline{H} &= \int \underline{H}(\underline{k}) e^{i\underline{k} \cdot \underline{r}} d\underline{k}. \end{aligned} \right\} \quad (1.4)$$

The time-dependent Fourier components $\underline{E}(\underline{k})$ and $\underline{H}(\underline{k})$ satisfy the following set of equations as a result of (1.3) (the dot indicates differentiation with respect to time)¹⁾

$$\left. \begin{aligned} \dot{\underline{H}}(\underline{k}) &= -i[\underline{k} \underline{E}(\underline{k})], \\ \dot{\underline{E}}(\underline{k}) &= i[\underline{k} \underline{H}(\underline{k})], \\ \underline{k} \underline{H}(\underline{k}) &= 0, \\ \underline{k} \underline{E}(\underline{k}) &= 0, \end{aligned} \right\} \quad (1.5)$$

to which we must add the reality condition of the field

$$\left. \begin{aligned} \underline{E}(-\underline{k}) &= \underline{E}^*(\underline{k}), \\ \underline{H}(-\underline{k}) &= \underline{H}^*(\underline{k}). \end{aligned} \right\} \quad (1.6)$$

¹⁾ We shall not consider the question of how the quantities $\underline{E}(\underline{k})$, $\underline{H}(\underline{k})$, etc., transform on going from one coordinate system to another. We note only that Equations (1.5) are equivalent to the relativistically invariant ones (1.3).

The two vectors $\underline{E}$ and $\underline{H}$ can be replaced by the vectors $\underline{E}$ and $\underline{f}$, eliminating $\underline{H}$ with the aid of (1.5), and obtaining

$$\underline{H}(\underline{k}) = i \left[\frac{\underline{k}}{k} \dot{\underline{E}}(\underline{k}) \right]. \quad (1.7)$$

Further, we may remove the necessity for a separate reality requirement by performing a substitution which leads automatically to satisfaction of Equations (1.6);²⁾

$$\left. \begin{aligned} \underline{E}(\underline{k}) &= N(\underline{k}) (f(\underline{k}) + f^*(-\underline{k})), \\ \dot{\underline{E}}(\underline{k}) &= -ik N(\underline{k}) (f(\underline{k}) - f^*(-\underline{k})). \end{aligned} \right\} \quad (1.8)$$

Here $N(\underline{k})$ is an arbitrary normalizing factor which, as we shall see below, should be chosen

$$N(\underline{k}) = \frac{\sqrt{k}}{4\pi^{3/2}}. \quad (1.9)$$

It is not difficult to obtain the equation satisfied by $f(\underline{k})$. Eliminating $\underline{H}$ from (1.5), we obtain

$$\left(\frac{\partial^2}{\partial t^2} + k^2 \right) \underline{E}(\underline{k}) = 0,$$

which can be rewritten in the form

$$\left(\frac{\partial}{\partial t} + ik \right) \left(\frac{\partial}{\partial t} - ik \right) \underline{E}(\underline{k}) = 0. \quad (1.10)$$

²⁾ This substitution actually consists of replacing the two real functions $\underline{E}$ and $\underline{H}$ by a single complex one

$$f(\underline{r}) = \int \underline{f}(\underline{k}) e^{i\underline{k} \cdot \underline{r}} d\underline{k}.$$

It should be born in mind, however, that $\underline{f}(\underline{r})$ cannot be represented in the form of a linear combination of $\underline{E}$ and $\underline{H}$. See in this respect, Section 2, paragraph 2.

Eliminating $\underline{f}^*(-\underline{k})$ from (1.8), we obtain

$$\left(\frac{\partial}{\partial t} - ik\right) E(\underline{k}) = -2iN(\underline{k}) \hbar f(\underline{k});$$

thus Equation (1.10) can be written

$$i \frac{\partial f(\underline{k})}{\partial t} = \hbar f(\underline{k}). \quad (1.11)$$

It is not difficult to see that $\underline{f}^*(\underline{k})$ satisfies the equation conjugate to (1.11). Equations (1.5), after expressions (1.7) and (1.8) have been inserted into them, give

$$\hbar f(\underline{k}) = 0. \quad (1.12)$$

Equations (1.11) and (1.12) are equivalent to Maxwell's equations. We note that (1.11) has the form of a Schrodinger equation if the energy operator $\mathbb{W}$ of the photon in $\underline{k}$ -space is defined as the operator of multiplication by the number $\underline{k}$, namely

$$\mathbb{W} = \underline{k} \quad (1.13)$$

Equation (1.13) is identical with (1.2). We shall show later that the function $\underline{f}(\underline{k})$ can be interpreted as the photon wave function in the usual quantum mechanical sense. We shall be able to define photon operators also for other physical quantities in $\underline{k}$ -space, for instance the momentum operator, the angular momentum operator, etc.

We note that not all solutions of the Schrodinger equation (1.11) correspond to actual photon states; only those solutions should be chosen which satisfy the transversality condition (1.12).

3. Energy.

We shall show that the operator $\mathbb{W}$ we have introduced can actually be considered the photon energy operator. Let us construct an expression for the energy of the electromagnetic field, which we shall designate $\underline{w}$ ¹⁾

$$\underline{w} = \frac{1}{2} \int (E^2 + H^2) d\mathbf{r}. \quad (1.14)$$

This is an integral over space of quantities which are quadratic in the field vectors. On the other hand, in quantum mechanics a space integral of expressions quadratic in the wave function is interpreted as the expectation value of the corresponding physical quantity. Therefore, the "corpuscular" interpretation of Equation (1.14) as the expectation value of the photon energy is a natural generalization.

Let us show that $\underline{w}$ can be written in the form

$$\underline{w} = \int f^*(\underline{k}) \mathbb{W} f(\underline{k}) d\mathbf{k}, \quad (1.15)$$

¹⁾ We shall take Heaviside units for $\underline{E}$ and $\underline{H}$.

where $\mathbb{W}$ is given by (1.13).

To do this let us insert the expansion (1.4) into (1.14), obtaining

$$\underline{w} = \frac{1}{2} \int (E(\underline{k}) E(\underline{k}') + H(\underline{k}) H(\underline{k}')) e^{i(\underline{k}+\underline{k}')\mathbf{r}} d\mathbf{k} d\mathbf{k}' d\mathbf{r}.$$

Carrying out the integration over $\mathbf{r}$, we arrive at

$$\int e^{i(\underline{k}+\underline{k}')\mathbf{r}} d\mathbf{r} = (2\pi)^3 \delta(\underline{k} + \underline{k}'),$$

where $\delta(\underline{k} + \underline{k}')$ is the three-dimensional Dirac δ -function; with (1.7), we obtain

$$\underline{w} = 4\pi^3 \int \left\{ E(\underline{k}) E(-\underline{k}) + \frac{1}{k^2} \dot{E}(\underline{k}) \dot{E}(-\underline{k}) \right\} d\mathbf{k}.$$

Finally, let us express $\underline{E}$ and $\dot{\underline{E}}$ in terms of $\underline{f}$ according to (1.8). This gives

$$\underline{w} = 8\pi^3 \int N^2 [f^*(\underline{k}) f(\underline{k}) + f^*(-\underline{k}) f(-\underline{k})] d\mathbf{k} = 16\pi^3 \int N^2 f^*(\underline{k}) f(\underline{k}) d\mathbf{k}.$$

If $\underline{N}(\underline{k})$ is chosen according to (1.9), we arrive at (1.15).

Let us consider the monochromatic solution of Equation (1.11):

$$f(\underline{k}) = f_0(\underline{k}) e^{-i\mathbf{p}t}, \quad (1.16)$$

where $\mathbf{p}$ is the eigenvalue of the photon energy operator

$$\mathbb{W} f_0(\underline{k}) = \mathbf{p} f_0(\underline{k}).$$

This shows that $\underline{f}(\underline{k})$ fails to vanish only for $\underline{k} = \underline{\mathbf{p}}$. The expression for the energy (1.15) then becomes

$$\underline{w} = \mathbf{p} \int f^*(\underline{k}) f(\underline{k}) d\mathbf{k}.$$

We require that this be identical with the fundamental quantum condition (1.2). Then the wave function $f(\mathbf{k})$ must satisfy the normalization condition

$$\int f^*(\mathbf{k}) f(\mathbf{k}) d\mathbf{k} = 1. \quad (1.17)$$

§ 2. Momentum Eigenstates.

1. Momentum

Let us now construct an expression for the momentum of the electromagnetic field, denoted by $\mathbf{p}$:

$$\mathbf{p} = \int \mathbf{E} \mathbf{H} d\mathbf{r}. \quad (2.1)$$

We shall interpret $\bar{\mathbf{p}}$ as the expectation value of the photon momentum.

Let us now express $\bar{\mathbf{p}}$ in terms of $f(\mathbf{k})$. Inserting the expansion (1.4) into (2.1), we obtain

$$\begin{aligned} \bar{\mathbf{p}} &= \int [\mathbf{E}(\mathbf{k}) \mathbf{H}(\mathbf{k}')] e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} d\mathbf{k} d\mathbf{k}' d\mathbf{r} = (2\pi)^3 \int [\mathbf{E}(\mathbf{k}) \mathbf{H}(-\mathbf{k})] d\mathbf{k} = \\ &= -(2\pi)^3 i \int \left[\mathbf{E}(\mathbf{k}) \left[\frac{\mathbf{k}}{k^2} \dot{\mathbf{E}}(-\mathbf{k}) \right] \right] d\mathbf{k} = -(2\pi)^3 i \int \frac{\mathbf{k}}{k^2} (\mathbf{E}(\mathbf{k}) \dot{\mathbf{E}}(-\mathbf{k})) d\mathbf{k}, \end{aligned}$$

or, writing $\mathbf{E}$ and $\dot{\mathbf{E}}$ according to (1.8) in terms of f ,

$$\bar{\mathbf{p}} = (2\pi)^3 \int N^2 \left\{ f^*(\mathbf{k}) f(\mathbf{k}) - f^*(-\mathbf{k}) f(-\mathbf{k}) + \right. \\ \left. + f^*(-\mathbf{k}) f^*(\mathbf{k}) - f(\mathbf{k}) f(-\mathbf{k}) \right\} \frac{\mathbf{k}}{k} d\mathbf{k}.$$

The last two terms vanish on integration, and each of the first two terms gives the same result. This is easy to show if we replace $\mathbf{k}$ by $-\mathbf{k}$ in the integrand. Making use of expression (1.9) for $N(\mathbf{k})$, we obtain

$$\bar{\mathbf{p}} = \int \mathbf{k} f_a^* f_a d\mathbf{k} \quad (2.2)$$

(the index $a = 1, 2, 3$ denotes the components of the vector $\mathbf{f}$; the summation convention is used).

It is thus valid to call the operator of multiplication by the wave vector

$$\mathbf{p} \equiv \mathbf{k} \quad (2.3)$$

the photon momentum operator, and to give $\mathbf{k}$ -space the name momentum space. The quantity $f^* f$ may be interpreted as the probability density that the photon possesses a momentum $\mathbf{k}$, and Equation (2.2) is then the usual quantum mechanical expression for the expectation value.

Thus, the normalization condition (1.17) has a simple and natural physical meaning.

2. On the Photon Wave Function in Configuration Space.

By performing an inverse Fourier transformation on $f(\mathbf{k})$,

$$\int f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{k} = f(\mathbf{r}),$$

we would be able to determine the photon wave function $f(\mathbf{r})$ in configuration space.¹⁾ In view of the normalization condition (1.17) for $f(\mathbf{k})$, $f(\mathbf{r})$ will also be normalized in the usual way, namely,

$$\int f^*(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} = 1.$$

However, the quantity $f^*(\mathbf{r}) f(\mathbf{r})$ cannot be interpreted as the probability density for finding the photon at a given space point. Indeed, the presence of a photon can be established only by its interaction with charges. This interaction is determined by the electromagnetic field vectors $\mathbf{E}$ and $\mathbf{H}$ at the given point. The latter, however, are not determined by the wave function $f(\mathbf{r})$ at that point, but by its values in all of space. This is due to the fact that the Fourier components of the field vectors [see (1.8)] expressed in terms of $f(\mathbf{k})$ contain the coefficient $\sqrt{k}$. Thus, the relation between $\mathbf{E}(\mathbf{r})$ and $f(\mathbf{r})$ will not be local, but is an integral relation.²⁾ In view of this situation, the localization of the photon in a region smaller than its wave length has no meaning, and the concept of probability density for a localized photon does not exist.

This result is strongly related to the behavior of particle densities under Lorentz transformations (see Section 53). It is impossible to construct bilinear combinations of the electromagnetic field vectors which are four-vectors satisfying the continuity equation³⁾ (although the energy-momentum tensor exists).

3. Plane Waves.

Let us return to a consideration of the photon wave function $f(\mathbf{k})$ in momentum space. Schrodinger's equation (1.11) determines its time dependence

$$f(\mathbf{k}, t) = f_0(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}t}.$$

¹⁾ L. Landau and R. Peierls, Z. Physik 62, 188 (1930); see also W. Pauli, Fundamental Principles of Wave Mechanics (State Tech. Press, 1947).

²⁾ Formally this can be written

$$\mathbf{E}(\mathbf{r}) = \sqrt{-\Delta} f(\mathbf{r}) + \text{complex conjugate},$$

where Δ is the Laplacian; $\sqrt{-\Delta}$ is actually an integral operator.

³⁾ This is a unique property that follows from the reality of the field (see Section 49).

The time-independent function $f_{\mathbf{p}}(\mathbf{k})$ is restricted only by the transversality condition (1.12). In particular, we may consider states for which f is different from zero only in the neighborhood of the point $\mathbf{k} = \mathbf{p}$ in the volume element $d\mathbf{p}$. These will be states with definite (actually almost definite) momentum, and the wave functions corresponding to them will be eigenfunctions¹⁾ of the momentum operator $\hat{\mathbf{p}}$. Bearing in mind the normalization condition (1.17), we can write the momentum eigenfunction in the form²⁾

$$f_{\mathbf{p}} = \frac{e_{\mathbf{p}}}{\sqrt{d\mathbf{p}}} e^{-i\mathbf{p} \cdot \mathbf{r}} \delta_{\mathbf{p}} \quad (2.4)$$

where $\mathbf{e}_{\mathbf{p}}$ is the photon polarization vector, of modulus unity and perpendicular to $\mathbf{p}$:

$$|\mathbf{e}_{\mathbf{p}}|^2 = 1, \quad \mathbf{e}_{\mathbf{p}} \cdot \mathbf{p} = 0.$$

This normalization corresponds to one photon in the volume element $d\mathbf{p} = dp_x dp_y dp_z$. In calculating the probabilities for various processes, a weighting factor is usually introduced (the number of states in the interval $d\mathbf{p}$). For the above normalization, this factor is equal to unity.

For a given $\mathbf{p}$, two linearly independent vectors $\mathbf{e}_{\mathbf{p}}$ are possible. Let the z axis be chosen along $\mathbf{p}$. These vectors can then be chosen in the following way:

$$\begin{matrix} e_{1x} = 1, & e_{1y} = 0, & e_{1z} = 0, \\ e_{2x} = 0, & e_{2y} = 1, & e_{2z} = 0 \end{matrix} \quad (2.5)$$

(linear polarization), or

$$\begin{matrix} e_{1x} = \frac{1}{\sqrt{2}}, & e_{1y} = \frac{i}{\sqrt{2}}, & e_{1z} = 0, \\ e_{2x} = \frac{1}{\sqrt{2}}, & e_{2y} = -\frac{i}{\sqrt{2}}, & e_{2z} = 0 \end{matrix} \quad (2.6)$$

(circular polarization). In both cases $\mathbf{e}_1$ and $\mathbf{e}_2$ are mutually orthogonal, so that

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = 0.$$

Thus, the momentum eigenstates are doubly degenerate. A unique characterization of a state requires also the knowledge of its polarization, and we have, therefore, provided the function (2.4) with the second index μ . The quantities p_x, p_y, p_z and μ are the complete set of photon quantum numbers (of course the energy is determined by these: $\omega = |\mathbf{p}|$). The set of functions $\frac{1}{\sqrt{d\mathbf{p}}}$ is a complete orthonormal set, and an arbitrary function $f(\mathbf{k})$ can be expanded in terms of these, writing

¹⁾ Actually proper differentials.

²⁾ The factor $\frac{1}{\sqrt{d\mathbf{p}}}$ in (2.4) has no profound meaning and is introduced only for agreement in the future with generally accepted notation (the wave function always contains an arbitrary phase factor).

$$\left. \begin{aligned} f(\mathbf{k}) &= \sum_{\mathbf{p}, \mu} C_{\mathbf{p}\mu} f_{\mathbf{p}\mu}, \\ C_{\mathbf{p}\mu} &= \int f f_{\mathbf{p}\mu}^* d\mathbf{k}. \end{aligned} \right\} \quad (2.7)$$

The expression for the electric and magnetic fields corresponding to a photon state with given momentum and polarization can, according to (1.4), (1.7) and (1.8) be written

$$\left. \begin{aligned} E_{\mathbf{p}\mu} &= \mathcal{E}_{\mathbf{p}\mu}(r) + \mathcal{E}_{\mathbf{p}\mu}^*(r), \\ H_{\mathbf{p}\mu} &= \mathcal{H}_{\mathbf{p}\mu}(r) + \mathcal{H}_{\mathbf{p}\mu}^*(r), \end{aligned} \right\} \quad (2.8)$$

where

$$\left. \begin{aligned} \mathcal{E}_{\mathbf{p}\mu} &= \frac{i}{4\pi^2} \sqrt{\frac{p}{d\mathbf{p}}} e_{\mu} e^{i(\mathbf{p} \cdot \mathbf{r} - \omega t)}, \\ \mathcal{H}_{\mathbf{p}\mu} &= \frac{i}{4\pi^2} \sqrt{\frac{p}{d\mathbf{p}}} \left[\frac{\mathbf{p}}{p} \cdot \mathbf{e}_{\mu} \right] e^{i(\mathbf{p} \cdot \mathbf{r} - \omega t)}. \end{aligned} \right\} \quad (2.9)$$

The normalization given in (2.9) corresponds to the existence in all space of a single photon whose momentum lies between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$. If, instead of the normalization (1.17), we were to choose the normalization

$$\int f_{\mathbf{p}\mu}^*(\mathbf{k}) f_{\mathbf{p}'\mu'}(\mathbf{k}) d\mathbf{k} = \delta(\mathbf{p} - \mathbf{p}') \delta_{\mu\mu'},$$

then the expressions for the fields would differ from (2.9) by the absence of the factor $\sqrt{\frac{p}{d\mathbf{p}}}$.

We shall in the future make use also of a normalization for which the photon is found in some large volume V , so that

$$\overline{\omega} = \frac{1}{V} \int (E^2 + H^2) d\mathbf{r} = \omega. \quad (2.10)$$

Then

$$\mathcal{E}_{\mathbf{p}\mu} = i \sqrt{\frac{\hbar}{2V}} e_{\mu} e^{i(\mathbf{p} \cdot \mathbf{r} - \omega t)} \quad (2.11)$$

and

$$3\mathcal{L}_{\mu\nu} = \left[\frac{p}{p'} \mathcal{E}_{\mu\nu} \right].$$

The decomposition (2.7) corresponds to decomposing an arbitrary electric field satisfying Maxwell's equations (1.3) into plane polarized waves:

$$E = \mathcal{E} + \mathcal{E}^*,$$

$$\mathcal{E} = \sum_{\mu=1}^3 C^\mu \mathcal{E}_{\mu},$$

where the coefficients C^μ are determined from (2.7). Analogous decompositions hold also for the magnetic field.

§ 3. Angular Momentum. Spin of the Photon.

1. Angular Momentum Operator.

Let us use the photon wave function $f(\mathbf{k})$ to express the angular momentum $\bar{\mathbf{M}}$ of the electromagnetic field, which we shall identify with the expectation value of the photon angular momentum in the state $f(\mathbf{k})$. The angular momentum of the field, as is well known, is given by the expression

$$\bar{\mathbf{M}} = \int [\mathbf{r} (E\mathbf{H})] d\mathbf{r}. \quad (3.1)$$

In terms of the Fourier transforms, we obtain

$$\bar{\mathbf{M}} = \int [\mathbf{r} (E(\mathbf{k}) H(\mathbf{k}'))] e^{i(\mathbf{k}+\mathbf{k}') \cdot \mathbf{r}} d\mathbf{k} d\mathbf{k}' d\mathbf{r}.$$

Let us first integrate over $\mathbf{r}$:

$$\int \mathbf{r} e^{i(\mathbf{k}+\mathbf{k}') \cdot \mathbf{r}} d\mathbf{r} = -i \nabla_{\mathbf{k}'} \int e^{i(\mathbf{k}+\mathbf{k}') \cdot \mathbf{r}} d\mathbf{r} = -i (2\pi)^3 \nabla_{\mathbf{k}'} \delta(\mathbf{k} + \mathbf{k}')$$

(here $\nabla_{\mathbf{k}'}$ means differentiation with respect to $\mathbf{k}'$). Now let us integrate over $\mathbf{k}'$ by parts:

$$\int d\mathbf{k}' [\nabla_{\mathbf{k}'} \delta(\mathbf{k} + \mathbf{k}') [E(\mathbf{k}) H(\mathbf{k}'))] = - \int d\mathbf{k}' \delta(\mathbf{k} + \mathbf{k}') [\nabla_{\mathbf{k}'} [E(\mathbf{k}) H(\mathbf{k}'))].$$

Replacing $\mathbf{H}$ by expression (1.7) which relates $\mathbf{H}$ to $\mathbf{E}$, simple operations lead to

$$[\nabla_{\mathbf{k}'} [E(\mathbf{k}) H(\mathbf{k}'))] =$$

$$= -i \left[\mathbf{k}' \nabla_{\mathbf{k}'} \left(\frac{E(\mathbf{k}) \dot{E}(\mathbf{k}')}{k'^2} \right) - i \mathbf{k}' E(\mathbf{k}) \text{curl}_{\mathbf{k}'} \frac{\dot{E}(\mathbf{k}')}{k'^2} + i \left[\frac{\dot{E}(\mathbf{k}')}{k'^2} E(\mathbf{k}) \right] \right].$$

Thus,

$$\bar{\mathbf{M}} = (2\pi)^3 \int d\mathbf{k} d\mathbf{k}' \delta(\mathbf{k} + \mathbf{k}') \left\{ \left[\mathbf{k}' \nabla_{\mathbf{k}'} \left(\frac{E(\mathbf{k}) \dot{E}(\mathbf{k}')}{k'^2} \right) \right] - \right.$$

$$\left. - \frac{1}{k'^2} [\dot{E}(\mathbf{k}') E(\mathbf{k})] + (\mathbf{k}' E(\mathbf{k})) \text{curl}_{\mathbf{k}'} \frac{\dot{E}(\mathbf{k}')}{k'^2} \right\}.$$

Let us now integrate over $\mathbf{k}'$. The last term vanishes, since

$$\mathbf{k} \mathbf{E}(\mathbf{k}) = 0,$$

and we obtain

$$\bar{\mathbf{M}} = (2\pi)^3 \int d\mathbf{k} \left\{ \left[\mathbf{k} \nabla \left(\frac{1}{k^2} E_c(\mathbf{k}) \dot{E}(-\mathbf{k}) \right) \right] - \frac{1}{k^2} [\dot{E}(-\mathbf{k}) E(\mathbf{k})] \right\}$$

(here ∇ refers to differentiation with respect to $\mathbf{k}$, and the index c on E_c indicates that in performing the differentiation this quantity is treated as a constant). Finally, we express $\dot{E}$ and $\ddot{E}$ in terms of f according to (1.8). This gives

$$\bar{\mathbf{M}} = \frac{i}{2} \int d\mathbf{k} \left\{ [\mathbf{k} \nabla (f_c(\mathbf{k}) f^*(-\mathbf{k}) - f_c(-\mathbf{k}) f(\mathbf{k})) + \right.$$

$$+ f_c^*(-\mathbf{k}) f^*(\mathbf{k}) - f_c(\mathbf{k}) f(-\mathbf{k})] -$$

$$- ([f^*(\mathbf{k}) f(\mathbf{k})] - [f(-\mathbf{k}) f^*(-\mathbf{k})]) +$$

$$+ [f^*(\mathbf{k}) f^*(-\mathbf{k})] - [f(-\mathbf{k}) f(\mathbf{k})] \}.$$

Terms containing both of the arguments $\mathbf{k}$ and $-\mathbf{k}$, vanish under integration. For instance,

$$\int [f(-\mathbf{k}) f(\mathbf{k})] d\mathbf{k},$$

when $\mathbf{k}$ is replaced by $-\mathbf{k}$, becomes

$$\int [f(k) f(-k)] dk,$$

which is the negative of the original expression. The integral

$$\int [k \nabla (f_c(k) f(-k))] dk$$

when $\underline{k}$ is replaced by $-\underline{k}$, becomes

$$\int [k \nabla (f_c(-k) f(k))] dk,$$

and, on the other hand, if we had integrated by parts we would obtain the same expression with the opposite sign. The integrals of the terms containing the argument $-\underline{k}$ do not change when $\underline{k}$ is replaced by $-\underline{k}$.

Again integrating by parts, we obtain

$$\int [k \nabla (f_c(k) f^*(k))] dk = - \int [k \nabla (f_c^*(k) f(k))] dk,$$

and arrive at the following expression for the angular momentum:

$$\bar{M} = \int dk \{ -i [k \nabla (f_c^* f)] - i [f^* f] \} dk,$$

or its components,

$$\bar{M}_\alpha = \int dk \{ f_\beta^* [-i k \nabla_\alpha f_\beta - i \epsilon_{\alpha\beta\gamma} f_\gamma] \}, \quad (3.2)$$

where $\epsilon_{\alpha\beta\gamma}$ is the antisymmetric unit tensor of rank three.

2. Spin Operator.

Let us introduce the vector operator $\underline{s}$, defined by its action on the components of $\underline{f}$ according to the equation

$$s_\alpha f_\beta = -i \epsilon_{\alpha\beta\gamma} f_\gamma. \quad (3.3)$$

We note that the vector product can be written in terms of $\underline{s}$ in the following way:

$$[E f]_\alpha = g (s_\alpha f). \quad (3.4)$$

On the basis of (3.3), we can write (3.2) in the form

$$\bar{M} = \int f_\alpha^* (-i [k \nabla] + s) f_\alpha dk. \quad (3.5)$$

We see that expression (3.5) for the expectation value of the photon angular momentum actually has the structure of a quantum mechanical expectation value with the angular momentum operator of the photon given by

$$M = -i [k \nabla] + s. \quad (3.6)$$

We shall show that (3.6) corresponds to the infinitesimal rotation operator of the vector field multiplied by $\hbar$. For this purpose let us consider an infinitesimal rotation about the origin in $\underline{k}$ -space.¹⁾ Such a rotation, as is well known, is defined by the infinitesimal rotation vector

$$\delta \alpha = \nu \delta \alpha,$$

where $\delta \alpha$ is the angle of rotation, and ν is the unit vector which defines the rotation axis. For such a rotation, the position vector of a point in space changes by an amount

$$\delta \underline{k} = [\delta \alpha \underline{k}].$$

An infinitesimal rotation in $\underline{k}$ -space transforms the vector field $\underline{f}(\underline{k})$ into the field $\underline{f}'(\underline{k})$ related to $\underline{f}(\underline{k})$ by the expression

$$f'(\underline{k} + \delta \underline{k}) = f(\underline{k}) + [\delta \alpha f(\underline{k})].$$

1) V. Sorokin, J. Exptl.-Theoret. Phys. 18, 228 (1948).

Then at a given point $\mathbf{k}$ the field changes by an amount

$$\delta f(\mathbf{k}) = f'(\mathbf{k}) - f(\mathbf{k}).$$

Since

$$f'(\mathbf{k} + \delta \mathbf{k}) = f'(\mathbf{k}) + (\delta \mathbf{k} \nabla) f'(\mathbf{k}) = f'(\mathbf{k}) + (\delta \mathbf{k} \nabla) f(\mathbf{k})$$

(in the last term the difference between $f'(\mathbf{k})$ and $f(\mathbf{k})$ may be neglected), we have

$$\delta f(\mathbf{k}) = -(\delta \mathbf{k} \nabla) f(\mathbf{k}) + [\delta \mathbf{k} f(\mathbf{k})]. \quad (3.7)$$

(We note that for a spherically symmetric field, which can always be written in the form $f(\mathbf{k}) = g(k)$, we have $\delta f = 0$.)

The infinitesimal rotation operator $\mathbf{J}$ is given by the expression

$$\delta f = (\delta \mathbf{a} \mathbf{J}) f = \delta \mathbf{a} (\mathbf{J} f).$$

Comparing this expression with (3.7) and using the definition of $\mathbf{s}$ given in (3.3) we find (note that $\mathbf{s}_{\alpha\beta} = -\mathbf{s}_{\beta\alpha}$, i.e., $\mathbf{s}(\mathbf{v} \mathbf{f}) = -(\mathbf{s} \mathbf{v}) \mathbf{f}$):

$$\mathbf{J} = -[\mathbf{k} \nabla] - \mathbf{s}.$$

Using expression (3.6) it is easy to show that the components of the angular momentum operator satisfy the following commutation rules:³⁾

$$\begin{aligned} M_x M_y - M_y M_x &= i M_z, \\ M_x M_z - M_z M_x &= 0. \end{aligned} \quad (3.8)$$

It follows from (3.8), as is well known, that the eigenvalues of M^2 are given by

$$M^2 = j(j+1), \quad (3.9)$$

where $2j+1$ is a positive integer. (We shall see below that j is also an integer.) The eigenvalues of the operator M_z are

$$M_z = M(M = -j, -j+1, \dots, j). \quad (3.10)$$

³⁾ These relations are valid for any quantum mechanical system and follow from the general connection between the angular momentum operators and the infinitesimal rotations.

The operator $\mathbf{M}$ commutes with the energy. Therefore, photon states with definite values of ω , M^2 and M_z may exist (the quantum numbers ω, j, M). We shall now concern ourselves with obtaining the wave functions of these states.

3. Spin Wave Functions.

Equation (3.6) shows that the photon angular momentum operator consists of two terms. The first term is the same as the usual quantum mechanical orbital angular momentum operator $\mathbf{L}$ in the momentum representation:

$$\mathbf{L} = -i[\mathbf{k} \nabla]. \quad (3.11)$$

The second term $\mathbf{s}$ may be called the spin angular momentum operator.

The separation of the photon angular momentum into its orbital and spin parts has a limited physical meaning. In the first place, the ordinary definition of spin as the angular momentum of a particle at rest is not applicable to the photon, since the photon rest mass is equal to zero. In the second place, states with definite values of the orbital and spin angular momenta, as we shall see below, do not in general satisfy the transversality condition. Therefore, only certain superpositions of these states have physical meaning. Nevertheless, the representation of the angular momentum in the form of two terms is extremely useful from the formal point of view. It allows us to construct wave functions for photon states with definite values of the angular momentum from the simpler eigenfunctions of the orbital and spin angular momenta.

The vector index α of the photon wave function may be considered an independent variable (it can be called the spin variable) which takes on three values: $\alpha = x, y, z$. Correspondingly, we shall introduce the notation

$$f_\alpha(\mathbf{k}) = f(\mathbf{k}, \alpha). \quad (3.12)$$

The function $f(\mathbf{k}, \alpha)$ is a scalar in the generalized spin and momentum space (the space of the variables k_x, k_y, k_z, α). Various components of the vector $\mathbf{f}_\alpha$ are now values of the scalar $f(\mathbf{k}, \alpha)$ at various points of spin space. The operator $\mathbf{L}$ operates only on the variable $\mathbf{k}$, and the operator $\mathbf{s}$ [see its definition (3.3)] operates only on α . Therefore, the operators $\mathbf{L}$ and $\mathbf{s}$ commute.

The eigenfunctions of the operators $\mathbf{L}^2$ and L_z , corresponding to eigenvalues

$$\left. \begin{aligned} L^2 &= l(l+1), \\ L_z &= m, \end{aligned} \right\} \quad (3.13)$$

shall be denoted by Φ_{lm} . They are functions only of the variables $\mathbf{k}$ and satisfy the equation

$$\left. \begin{aligned} L^2 \Phi_{lm} &= l(l+1) \Phi_{lm}, \\ L_z \Phi_{lm} &= m \Phi_{lm}. \end{aligned} \right\} \quad (3.14)$$

The solutions of (3.14), as is very well known, are the spherical functions

$$\Phi_{lm}(\mathbf{k}) = a(k) Y_{lm}(\pi), \quad (3.15)$$

where $\pi = \frac{\mathbf{k}}{k}$. We shall use the spherical functions normalized according to

$$\int Y_{lm} Y_{l'm'} d\Omega = \delta_{ll'} \delta_{mm'}. \quad (3.16)$$

Thus, the functions $\chi_{\pm\mu}(\mathbf{k})$ should be normalized so that

$$\int_0^\infty a^*(k) a(k) k^3 dk = 1. \quad (3.17)$$

Let us now find the eigenfunctions of the spin operator. Let $\chi_{\pm\mu}$ be an eigenfunction corresponding to the following eigenvalues of $\hat{s}^2$ and $\hat{s}_z$:

$$\left. \begin{aligned} s^2 &= s(s+1), \\ s_z &= \mu. \end{aligned} \right\} \quad (3.18)$$

The argument of $\chi_{\pm\mu}(\alpha)$ is the spin variable α . We may thus write this function in the form of a vector $\chi_{\pm\mu}$.

If χ is written as a column vector

$$\chi = \begin{pmatrix} \chi_x \\ \chi_y \\ \chi_z \end{pmatrix}$$

then Equation (3.3) can be used to obtain the operators s_x and s^2 in the form of the following matrices:

$$\left. \begin{aligned} s_x &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}; & s_y &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}; \\ s_z &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; & s^2 &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned} \right\} \quad (3.19)$$

From (3.19) it is seen that the quantity μ in (3.18) can take on only the value $\mu=1$. In other words, the photon spin is equal to one. We shall therefore sometimes suppress the index μ in the function $\chi_{\pm\mu}$, writing

$$\chi_{\pm\mu} = \chi_{\pm}.$$

The $\pm$ component μ of the spin can take on three values:

$$\mu = 0, \pm 1.$$

The functions $\chi_{\pm\mu}$ satisfy the equations

$$\begin{aligned} s^2 \chi_{\pm\mu} &= s(s+1) \chi_{\pm\mu}, \\ s_z \chi_{\pm\mu} &= \mu \chi_{\pm\mu}. \end{aligned}$$

The solution to these equations can be found on the basis of (3.19):

$$\chi_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \quad \chi_1 = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}; \quad \chi_{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}. \quad (3.20)$$

The functions (3.20) are mutually orthogonal, since they are eigenfunctions of the Hermitian operator s_z belonging to different eigenvalues. They are normalized so that

$$\sum_{\mu} \chi_{\mu}^*(\alpha) \chi_{\mu'}(\alpha) = \delta_{\mu\mu'},$$

or in vector form

$$\chi_{\mu}^* \chi_{\mu'} = \delta_{\mu\mu'}. \quad (3.21)$$

The unit vectors χ_{μ} define a basis in terms of which an arbitrary vector $\underline{f}$ can be resolved;

$$\underline{f} = \sum_{\mu=-1}^1 f^{\mu} \chi_{\mu}. \quad (3.22)$$

We shall call f^{μ} the contravariant components of the vector $\underline{f}$ in the coordinate system defined by this basis. With the expression χ_{μ} given in (3.20), it is easy to establish the relation between the f^{μ} and the cartesian components of the vector $\underline{f}$ (f_x, f_y, f_z):

$$\left. \begin{aligned} f^0 &= f_z, \\ f^{\pm 1} &= \mp \frac{1}{\sqrt{2}} (f_x \mp i f_y). \end{aligned} \right\} \quad (3.23)$$

In addition to the contravariant components f^{μ} , we shall make use of the covariant components f_{μ} , which are defined by the requirement that the scalar product of two vectors $\underline{f}$ and $\underline{g}$:

$$fg = f_x g_x + \frac{1}{2} (f_y + if_z) (g_x - ig_y) + \frac{1}{2} (f_y - if_z) (g_x + ig_y)$$

have the following form:

$$fg = \sum_{\mu=-1}^1 g^{\mu} f_{\mu} = \sum_{\mu=-1}^1 f^{\mu} g_{\mu}$$

This will be true, as is easy to see, if

$$f_{\mu} = (-1)^{\mu} f^{-\mu}. \quad (3.24)$$

We note that since the spin operator commutes with the momentum operator, we may speak of eigenstates of both the momentum and the projection of the spin. The components of the polarization vector in (2.4) may be chosen so that

$$\mathbf{e}_{\mu} = \chi_{\mu}.$$

The two possible polarizations correspond exactly to the two values of the spin projection μ . The third value is excluded by the transversality condition. If the z axis is directed along $\mathbf{p}$, the transversality condition excludes the state χ_0 . The two polarization vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ in (2.6) are equivalent to χ_1 and χ_{-1} , respectively. Thus, the value $\mu=1$ corresponds to right circular polarization, and $\mu=-1$ to left circular polarization.

§ 4. Angular Momentum and Parity Eigenstates.

1. Angular Momentum Eigenfunction.

Let us go on to a consideration of the eigenfunctions $f_{\mathbf{M}}$ of the operators $\mathbf{M}^2$ and M_z . These functions satisfy the equations:

$$\left. \begin{aligned} M^2 f_{\mathbf{M}}(k, a) &= j(j+1) f_{\mathbf{M}}(k, a), \\ M_z f_{\mathbf{M}}(k, a) &= M f_{\mathbf{M}}(k, a). \end{aligned} \right\} \quad (4.1)$$

Instead of solving these equations directly, we may make use of the quantum mechanical rules for composition of angular momenta in order to determine the functions $f_{\mathbf{M}}$. Indeed, our problem reduces to the well-known quantum mechanical problem of constructing wave functions of a system consisting of two noninteracting subsystems. In our case the subsystems are the orbital (variables $\mathbf{k}$) and spin (variable a) degrees of freedom of the photon.

Since the spin angular momentum $s=1$, then according to the rules for composition of angular momentum, the total angular momentum of the photon can take on a value j if the orbital angular momentum l is given by

$$l = j, j \pm 1 \quad (j \neq 0). \quad (4.2)$$

Thus, in the general case there exist three different wave functions $f_{\mathbf{M}}$ corresponding to three orbital states. We shall denote them by $f_{j\mathbf{M}}$.

The wave function $f_{j\mathbf{M}}$ is, as is well known, a superposition of products of orbital and spin functions $\Phi_{lm} \chi_s \mu$, in which the projections m and μ are related by the rule

$$M = m + \mu.$$

Therefore,

$$f_{j\mathbf{M}}(k, a) = \sum_{m+\mu=M} C_{jM}^{lm\mu} \Phi_{lm}(k) \chi_{s\mu}(a). \quad (4.3)$$

The general expressions for the coefficients $C_{jM}^{lm\mu}$ are known from elementary quantum mechanics.⁹ For our case ($s=1$) their values will be given below [see (4.3)].

Let us rewrite (4.3) in vector form:

$$f_{j\mathbf{M}}(k) = a(k) \sum_{\mu=-1}^1 C_{jM}^{M-\mu\mu} Y_{l, M-\mu}(n) \chi_{\mu}.$$

The orbital function Φ_{lm} is expressed in terms of the spherical function Y_{lm} according to (3.15). Therefore, according to the general resolution (3.22), the contravariant components of the vector $f_{j\mathbf{M}}$ are

$$f_{jM}^{\alpha} = a C_{jM}^{M-\mu\mu} Y_{l, M-\mu}^{\alpha}. \quad (4.4)$$

Correspondingly, the covariant components are

$$(f_{j\mathbf{M}})_{\alpha} = a C_{jM}^{M-\mu\mu} Y_{l, M-\mu, \alpha}, \quad (4.5)$$

where

$$C_{jM}^{M-\mu\mu} = (-1)^{\mu} C_{jM}^{j, M+\mu; \mu, -\mu}. \quad (4.6)$$

For given j and M , the coefficients $C_{jM}^{M-\mu\mu}$ form a matrix of three columns ($j=1, 2, 3$) and three rows ($\mu=0, \pm 1$). This matrix is orthogonal, i.e.,

⁹ See, for instance, L. Landau and E. Lifshitz, Quantum Mechanics (State Tech. Press, 1948); E. Condon and G. Shortley, The Theory of Atomic Spectra (Foreign Lit. Press, 1949).

⁹ V. Berestetsky, J. Exptl.-Theoret. Phys. 17, 12 (1947).

$$fg = f_x g_x + \frac{1}{2} (f_y + if_z) (g_y - ig_z) + \frac{1}{2} (f_y - if_z) (g_y + ig_z)$$

have the following form:

$$fg = \sum_{\mu=-1}^1 g^{\mu} f_{\mu} = \sum_{\mu=-1}^1 f^{\mu} g_{\mu}$$

This will be true, as is easy to see, if

$$f_{\mu} = (-1)^{\mu} f^{-\mu}. \quad (3.24)$$

We note that since the spin operator commutes with the momentum operator, we may speak of eigenstates of both the momentum and the projection of the spin. The components of the polarization vector in (2.4) may be chosen so that

$$\mathbf{e}_{\mu} = \chi_{\mu}$$

The two possible polarizations correspond exactly to the two values of the spin projection μ . The third value is excluded by the transversality condition. If the z axis is directed along $\mathbf{p}$, the transversality condition excludes the state χ_0 . The two polarization vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ in (2.6) are equivalent to χ_1 and χ_{-1} , respectively. Thus, the value $\mu=1$ corresponds to right circular polarization, and $\mu=-1$ to left circular polarization.

§ 4. Angular Momentum and Parity Eigenstates.

1. Angular Momentum Eigenfunction.

Let us go on to a consideration of the eigenfunctions f_{JM} of the operators $\mathbf{M}^2$ and M_z . These functions satisfy the equations:

$$\left. \begin{aligned} M^2 f_{JM}(k, a) &= j(j+1) f_{JM}(k, a), \\ M_z f_{JM}(k, a) &= M f_{JM}(k, a). \end{aligned} \right\} \quad (4.1)$$

Instead of solving these equations directly, we may make use of the quantum mechanical rules for composition of angular momenta in order to determine the functions f_{JM} . Indeed, our problem reduces to the well-known quantum mechanical problem of constructing wave functions of a system consisting of two noninteracting subsystems. In our case the subsystems are the orbital (variables $\mathbf{k}$) and spin (variable a) degrees of freedom of the photon.

Since the spin angular momentum $s=1$, then according to the rules for composition of angular momentum, the total angular momentum of the photon can take on a value J if the orbital angular momentum L is given by

$$L = J, J \pm 1 \quad (J \neq 0). \quad (4.2)$$

Thus, in the general case there exist three different wave functions f_{JM} corresponding to three orbital states. We shall denote them by f_{JM} .

The wave function f_{JM} is, as is well known, a superposition of products of orbital and spin functions $\phi_{JM} \chi_{SM}$, in which the projections m and μ are related by the rule

$$M = m + \mu.$$

Therefore,

$$f_{JM}(k, a) = \sum_{m+\mu=M} C_{JM}^{j, m; \mu} \phi_{jm}(k) \chi_{\mu\mu}(a). \quad (4.3)$$

The general expressions for the coefficients $C_{JM}^{j, m; \mu}$ are known from elementary quantum mechanics.⁹ For our case ($j=1$) their values will be given below [see (4.8)].

Let us rewrite (4.3) in vector form:

$$f_{JM}(k) = a(k) \sum_{\mu=-1}^1 C_{JM}^{j, M-\mu; \mu} Y_{1, M-\mu}(n) \chi_{\mu}.$$

The orbital function ϕ_{JM} is expressed in terms of the spherical function Y_{JM} according to (3.15). Therefore, according to the general resolution (3.22), the contravariant components of the vector f_{JM} are

$$f_{JM}^{\mu} = a C_{JM}^{j, M-\mu; \mu} Y_{1, M-\mu}. \quad (4.4)$$

Correspondingly, the covariant components are

$$(f_{JM})_{\mu} = a \epsilon_{\mu}^{j, M-\mu; \mu} Y_{1, M-\mu}, \quad (4.5)$$

where

$$\epsilon_{\mu}^{j, M-\mu; \mu} = (-1)^{\mu} C_{JM}^{j, M-\mu; \mu}. \quad (4.6)$$

For given J and M , the coefficients $C_{JM}^{j, M-\mu; \mu}$ form a matrix of three columns ($j=1, 2, 3$) and three rows ($\mu=0, \pm 1$). This matrix is orthogonal, i.e.,

⁹ See, for instance, L. Landau and E. Lifshitz, Quantum Mechanics (State Tech. Press, 1948); E. Condon and G. Shortley, The Theory of Atomic Spectra (Foreign Lit. Press, 1949).

⁹ V. Berestetsky, J. Exptl.-Theoret. Phys. 17, 12 (1947).

$$\left. \begin{aligned} \sum_{\mu} C_{JM}^{L, M-\mu, \mu} C_{JM}^{L, M-\mu, \mu} &= \delta_{JM}, \\ \sum_{\mu} C_{JM}^{L, M-\mu, \mu} C_{JM}^{L, M-\mu, \mu'} &= \delta_{\mu\mu'}. \end{aligned} \right\} \quad (4.7)$$

The explicit expression for the matrix $\underline{C}$ is given here:

"The Matrix Element" $C_{JM}^{L, M-\mu, \mu}$ ($s=1$)

μ	$l+1$	l	$l-1$
-1	$\sqrt{\frac{(l-M)(l-M+1)}{(2l+1)(2l+2)}}$	$\sqrt{\frac{(l+M+1)(l-M)}{2l(2l+1)}}$	$\sqrt{\frac{(l+M)(l+M+1)}{2l(2l+1)}}$
0	$\sqrt{\frac{(l+M+1)(l-M+1)}{(2l+1)(l+1)}}$	$\frac{M}{\sqrt{l(l+1)}}$	$-\sqrt{\frac{(l+M)(l-M)}{l(2l+1)}}$
1	$\sqrt{\frac{(l+M)(l+M+1)}{(2l+1)(2l+2)}}$	$-\sqrt{\frac{(l+M)(l-M+1)}{2l(2l+1)}}$	$\sqrt{\frac{(l-M)(l-M+1)}{2l(2l+1)}}$

(4.8)

Since the f_{JM} are an orthogonal set of functions (if any of the indices of two functions are different, that means that these functions belong to different eigenvalues of the Hermitian operators $\mathbf{M}^2$, M_z , and $\mathbf{L}^2$), the normalization conditions (4.7) and (3.17) lead to the relation

$$\int f_{JM} f_{J'M'} d\mathbf{k} = \delta_{JJ'} \delta_{MM'}. \quad (4.9)$$

We shall call the angular part of the function f_{JM} the vector spherical function or spherical vector and denote it by $\mathbf{Y}_{JM}$ ¹⁾. Equations (4.4) or (4.5) will be written in the form

$$f_{JM} = a(k) Y_{JM}(\mathbf{n}), \quad (4.10)$$

where

$$\left. \begin{aligned} (Y_{JM})_0 &= s_{JM}^{L, J-1} Y_{L, M+1}, \\ (Y_{JM})^\mu &= C_{JM}^{L, M-\mu, \mu} Y_{L, M-\mu}. \end{aligned} \right\} \quad (4.11)$$

Let us bear in mind that according to definitions (3.23) and (3.24), the covariant components of the vector $\mathbf{Y}$, namely the Y_μ , are given in terms of its cartesian components by the expression

¹⁾ For a definition of spherical vectors, see G. Petrashev, Proc. Acad. Sci., USSR 46, 291 (1945); V. Sorokin, J. Exptl.-Theoret. Phys. 18, 228 (1948); V. Berestetsky, J. Exptl.-Theoret. Phys. 17 (1947); V. Berestetsky, A. Dolginov, and K. Ter-Martirosyan, J. Exptl.-Theoret. Phys. 20, 527 (1950).

$$Y_0 = Y_z, \quad Y_{\pm 1} = \mp \frac{1}{\sqrt{2}} (Y_x \pm i Y_y).$$

From (4.9) and (3.17) it follows that the vector spherical functions are an orthonormal set:

$$\int Y_{JM} Y_{J'M'} d\mathbf{n} = \delta_{JJ'} \delta_{MM'}. \quad (4.12)$$

2. Spherical Vectors. Parity.

We have found a set of eigenfunctions of the operators corresponding to the square of the photon angular momentum $\mathbf{M}^2$ and to its projection M_z . A photon state with definite values of $\mathbf{J}$ and $\mathbf{M}$ is described by a wave function which is in general a linear combination of three spherical waves, namely

$$f_{JM} = \sum_{\lambda=-1}^{+1} c_\lambda Y_{JM}.$$

The coefficients of this linear combination are not independent, since the photon wave function must satisfy the transversality condition (1.12):

$$f_{JM} n = 0.$$

Therefore, there are not three, but two different photon states with given quantum numbers $\mathbf{J}$ and $\mathbf{M}$. The corresponding wave functions shall be denoted by $f_{JM}^{(\lambda)}$, where λ may take on the values $\lambda=1, 0$.

In order to obtain an explicit expression for $f_{JM}^{(\lambda)}$, we note that from the three linearly independent vectors $\mathbf{X}_{JM}$ in $\mathbf{k}$ -space, we can construct three linear combinations $\mathbf{X}_{JM}^{(\lambda)}$, with $\lambda=0, \pm 1$ (we shall call these also spherical vectors) such that they are mutually perpendicular, one of these say $\mathbf{X}_{JM}^{(-1)}$ can be made longitudinal, i.e., directed along the radius vector $\mathbf{k}$, and the two others ($\mathbf{X}_{JM}^{(1)}$ and $\mathbf{X}_{JM}^{(0)}$) transverse. Let us find these linear combinations.

Let us make use of the well known formula for the expansion of a product $\mathbf{n}_\mu X_{JM}$ in spherical functions, where $\mathbf{n}_\mu$ ($\mu=0, \pm 1$) are the components of a unit vector expressed in terms of the spherical coordinates ϑ and φ by the expression

$$n_0 = \cos \vartheta; \quad n_{\pm 1} = \mp \frac{1}{\sqrt{2}} \sin \vartheta e^{\pm i\varphi},$$

and X_{JM} is a spherical function.²⁾ In our notation this expansion can be written

²⁾ See, for instance, H. Bethe, Quantum Mechanics of Simple Systems (United Sci. Tech. Press, 1935), p. 383. [This is a translation of "Quantenmechanik der Ein und Zwei Elektronenprobleme," Handbuch der Physik, 2nd Ed., XXIV, Part 1 (1933).]

$$nY_{JM} = \sqrt{\frac{j}{2j+1}} Y_{j,j-1,M} - \sqrt{\frac{j+1}{2j+1}} Y_{j,j+1,M}, \quad (4.13)$$

which is easily verified from the definitions (4.11) for the spherical vectors and (4.8) for the $\epsilon_{\frac{j+1}{2}, \frac{j-1}{2}}^{\frac{j+1}{2}, \frac{j-1}{2}}$ as well as (4.8). Thus, Equation (4.13) is that linear combination of spherical vectors which gives a longitudinal vector. We therefore define

$$Y_{JM}^{(1)} = nY_{JM}. \quad (4.14)$$

Further, the scalar product of the unit vector $\hat{n}$ and the spherical vector $\mathbf{Y}_{JM}$

$$nY_{JM} = \sum_{\mu} (-1)^{\mu} \delta_{\mu, n-\mu} Y_{j, \mu, M+\mu}$$

vanishes, which can be shown by using (4.13) and (4.8). Thus, $\mathbf{Y}_{JM}$ is one of the desired transverse spherical vectors. Assigning the value $\lambda = 0$ to this one, we have

$$Y_{JM}^{(0)} = Y_{JM}. \quad (4.15)$$

Finally, the second transverse spherical vector $\mathbf{Y}_{JM}^{(2)}$ will be defined by the equation

$$iY_{JM}^{(2)} = [nY_{JM}^{(1)}]. \quad (4.16)$$

Using the expansion for nY_{JM} , we can express $\mathbf{Y}_{JM}^{(1)}$ in terms of the spherical vectors $\mathbf{Y}_{j, \mu, M}$, namely

$$Y_{JM}^{(1)} = \sqrt{\frac{j}{2j+1}} Y_{j, j+1, M} - \sqrt{\frac{j+1}{2j+1}} Y_{j, j-1, M}. \quad (4.17)$$

From definitions (4.14), (4.15), and (4.16) it can be seen that the vector spherical functions $\mathbf{Y}_{JM}^{(\lambda)}$ are normalized in the same way as the $\mathbf{Y}_{JM}$. The functions $\mathbf{Y}_{JM}^{(\lambda)}$, as do the $\mathbf{Y}_{JM}$, remain mutually orthogonal for different j and M . Since, furthermore, they are mutually perpendicular for different λ at every point, we obtain

$$\int Y_{JM}^{(\lambda)} Y_{JM'}^{(\lambda')} d\Omega = \delta_{JJ'} \delta_{MM'} \delta_{\lambda\lambda'}. \quad (4.18)$$

The wave function of a photon in an angular momentum eigenstate can be written

$$f_{JM} = \rho_1 Y_{JM}^{(1)} + \rho_0 Y_{JM}^{(0)}, \quad (4.19)$$

where the coefficients ρ_0 and ρ_1 are arbitrary. This means that the state $\mathbf{f}_{JM}$ is doubly degenerate. We can remove this degeneracy by requiring that each state also have a definite parity, i.e., be an eigenstate of the inversion operator I . For a vector field this is expressed by ¹⁾

$$If(\mathbf{k}) = -f(-\mathbf{k}).$$

The operator I commutes with the angular momentum and has two eigenvalues, equal to ± 1 ($I^2 = 1$). Since the values of a spherical function at the points $\mathbf{k}$ and $-\mathbf{k}$ are related by

$$Y_{lm}(-\mathbf{n}) = (-1)^l Y_{lm}(\mathbf{n}),$$

the definition (4.11) of the spherical vectors gives

$$IY_{JM}(\mathbf{n}) = (-1)^{J+1} Y_{JM}(\mathbf{n}).$$

For the transverse spherical vector, as can be seen from (4.15) and (4.17), we have

$$\left. \begin{aligned} IY_{JM}^{(1)} &= (-1)^J Y_{JM}^{(1)}, \\ IY_{JM}^{(0)} &= (-1)^{J+1} Y_{JM}^{(0)}. \end{aligned} \right\} \quad (4.20)$$

Thus, for a given j and M there exist two possible states differing in parity. The wave functions of these states shall be denoted by $\mathbf{f}_{JM\lambda}$ ($\lambda = 0, 1$).

The state with $\lambda = 1$ is called an electric state, and that with $\lambda = 0$ a magnetic state. These names are related to the fact that emission of a photon in the corresponding states is determined, as we shall see below (see Section 30), by the electric or magnetic moment of the system of charges.

If, in addition to the angular momentum and parity, the energy of the photon is definite (or almost definite), then a $(\mathbf{k})$ differs from zero in a small region $d\Omega$ in the neighborhood of $\mathbf{k} = \mathbf{p}$, that is $\mathbf{a}(\mathbf{k}) = \mathbf{a}_{\mathbf{p}} \delta_{\mathbf{k}\mathbf{p}}$, where, according to the normalization condition (1.17), ²⁾

$$a_{\mathbf{p}} = \frac{1e^{-i\mathbf{p}\cdot\mathbf{r}}}{\sqrt{p^3 d\mathbf{p}}}.$$

Thus, the photon state may be uniquely characterized by the four quantum numbers corresponding to the energy p , the angular momentum J , the projection of the angular momentum M , and the parity λ . Its normalized wave function can be written

$$f_{JM\lambda} = \frac{1}{\sqrt{p^3 d\mathbf{p}}} Y_{JM}^{(\lambda)} e^{-i\mathbf{p}\cdot\mathbf{r}} \delta_{\mathbf{k}\mathbf{p}}. \quad (4.21)$$

¹⁾ The minus sign arises because of the change of direction due to inversion.

²⁾ As in Equation (2.4), the factor i has no fundamental significance here.

We note that when $\lambda=1$, the photon cannot be assigned a definite value of L , since according to (4.17) the spherical vector $\mathbf{Y}_{JM}^{(1)}$ is a linear combination of spherical vectors $\mathbf{Y}_{JM}$ with different values of L . This is a manifestation of the fact that it is in reality impossible to divide the angular momentum of the photon into an orbital and spin part, as has been previously mentioned.

For the case $j=0$ there exists, according to Equation (4.2), only one spherical vector $\mathbf{Y}_{00}$. It is easy to see that it is a longitudinal one, since according to (4.14) the longitudinal spherical vector

$$\mathbf{Y}_{00}^{(-)} = n\mathbf{Y}_{00}$$

can always be constructed. It follows then that transverse spherical vectors do not exist for $j=0$. This result has a simple meaning. The state with zero angular momentum is a spherically symmetric one, but a spherically symmetric vector field can only be a longitudinal one.

Thus, the photon cannot exist in a state with angular momentum zero.

3. Expansion in Spherical Waves.

The wave function of the photon in an arbitrary state $\underline{f}$ can be expanded in a series of the functions given by (4.21):

$$\underline{f} = \sum_{JM\lambda} C_{JM\lambda}^{f} f_{JM\lambda}. \quad (4.22)$$

In view of the orthonormality of the set (4.21), the expansion coefficients are given by

$$C_{JM\lambda}^{f} = \int \underline{f} f_{JM\lambda}^* dk. \quad (4.23)$$

We can make use of expressions (4.22) and (2.7) to expand a state of definite momentum in a series of angular momentum eigenfunctions, and vice versa:

$$f_{\mathbf{p}} = \sum_{JM\lambda} C_{JM\lambda}^{f_{\mathbf{p}}} f_{JM\lambda}, \\ f_{JM\lambda} = \sum_{\mathbf{p}} C_{JM\lambda}^{f_{\mathbf{p}}} f_{\mathbf{p}},$$

where according to (2.4) and (4.21) ($d\mathbf{p} = p^2 dp d\Omega_p$)

$$C_{JM\lambda}^{f_{\mathbf{p}}} = (C_{JM\lambda}^{f_{\mathbf{p}}})^* = e_{\lambda} Y_{JM}^{(\lambda)}\left(\frac{\mathbf{p}}{p}\right) \sqrt{d\Omega_p} f_{\mathbf{p}}. \quad (4.24)$$

The magnitude of $\left| C_{JM\lambda}^{f_{\mathbf{p}}} \right|^2$ determines the probability that the photon is moving in the direction $\mathbf{p}$ and has a given polarization λ , if it is known that it has definite angular momentum and parity. Summing over polarization states, we obtain from (4.24)

$$\sum_{\lambda} |C_{JM\lambda}^{f_{\mathbf{p}}}|^2 = \left| Y_{JM}^{(0)}\left(\frac{\mathbf{p}}{p}\right) \right|^2 d\Omega_p. \quad (4.25)$$

We note that in view of the relation between two spherical vectors, as given by Equation (4.16), expression (4.25) does not depend on λ ; this means that the angular distribution of the photon is determined only by its angular momentum, and not by its parity. Here we present explicit expressions for the functions (4.25).

4. Expressions for the Electric and Magnetic Fields.

Let us find the electric and magnetic fields corresponding to the energy, angular momentum, and parity eigenstates of the photon. According to (1.8) and (4.21), the Fourier components of the electric field are given by

$$E_{JM\lambda}(k) = -\frac{i}{4\pi^2 \sqrt{p} dp} (Y_{JM}^{(\lambda)} e^{-ip\tau} - Y_{JM}^{(\lambda)*} e^{ip\tau}) \quad (\lambda = 0, 1). \quad (4.26)$$

Correspondingly, the Fourier components of the magnetic field are given by

$$H_{JM\lambda} = -\frac{1}{4\pi^2 \sqrt{p} dp} (Y_{JM}^{(\lambda-1)} e^{-ip\tau} + Y_{JM}^{(\lambda-1)*} e^{ip\tau}). \quad (4.27)$$

When (4.26) and (4.27) are inserted in the Fourier series (1.4), there occur integrals of the type

$$\int Y_{JM}^{(\lambda)}\left(\frac{\mathbf{k}}{k}\right) e^{ikr} d\Omega_k.$$

In order to calculate these, we shall use the well-known expansion of a plane wave in spherical functions:

$$e^{ikr} = \sum_{lm} g_l(kr) Y_{lm}^*(\frac{\mathbf{k}}{k}) Y_{lm}\left(\frac{\mathbf{r}}{r}\right), \quad (4.28)$$

where the radial function $g_l(kr)$ is given by

$$g_l(kr) = (2\pi)^{1/2} i^l \frac{J_{l+1/2}(kr)}{\sqrt{kr}} \quad (4.29)$$

[$J_{l+1/2}(z)$ is a Bessel function]. The function g_l satisfies the following normalization condition, which is not difficult to obtain from the asymptotic expansion of the Bessel function:

The Functions $|Y_{JM}^{(0)}(\theta)|^2$
($x = \cos \theta$)

$J \backslash M$	1	2	3	4	5
0	$-\frac{3}{8\pi}(x^2-1)$	$-\frac{15}{8\pi}(x^4-x^2)$	$-\frac{21}{64\pi}(25x^6-35x^4+11x^2-1)$	$-\frac{45}{64\pi}(49x^8-91x^6+51x^4-9x^2)$	$-\frac{165}{312\pi}(441x^{10}-1029x^8+885x^6-266x^4+29x^2-1)$
1	$\frac{3}{16\pi}(x^2+1)$	$\frac{5}{16\pi}(4x^4-3x^2+1)$	$\frac{7}{256\pi}(235x^6-305x^4+111x^2+1)$	$\frac{9}{256\pi}(784x^8-1463x^6+855x^4-153x^2+9)$	$\frac{11}{1024\pi}(11025x^{10}-26019x^8+21375x^6-7070x^4+813x^2+1)$
2		$-\frac{5}{16\pi}(x^4-1)$	$-\frac{35}{128\pi}(9x^6-11x^4+3x^2-1)$	$-\frac{9}{128\pi}(196x^8-371x^6+225x^4-49x^2-1)$	$-\frac{77}{256\pi}(225x^{10}-549x^8+474x^6-170x^4+21x^2-1)$
3			$\frac{105}{256\pi}(x^6-x^4-x^2+1)$	$\frac{63}{256\pi}(16x^8-31x^6+15x^4-x^2+1)$	$\frac{231}{2048\pi}(225x^{10}-579x^8+190x^6-29x^4+1)$
4				$-\frac{63}{128\pi}(x^8-2x^6+2x^4-1)$	$-\frac{231}{1024\pi}(25x^{10}-69x^8+10x^6-3x^4-1)$
5					$\frac{1155}{2048\pi}(x^{10}-3x^8+2x^6+2x^4-3x^2+1)$

$$\int_0^\infty g_1(kr) g_1^*(k'r) r^3 dr = (2\pi)^3 \frac{1}{k^3} \delta(k-k').$$

Using the explicit expressions for the components of the wave vector (4.11), we obtain

$$\int Y_{JM}^{(0)}\left(\frac{k}{k}\right) e^{ikr} d\Omega_k = g_1(kr) Y_{JM}\left(\frac{r}{r}\right).$$

According to the definition of the transverse wave vectors, it follows from this that

$$\begin{aligned} \int Y_{JM}^{(0)}\left(\frac{k}{k}\right) e^{ikr} d\Omega_k &= g_1(kr) Y_{JM}^{(0)}\left(\frac{r}{r}\right), \\ &= \sqrt{\frac{J}{2J+1}} g_{J+1}(kr) Y_{J,J+1,M}\left(\frac{r}{r}\right) + \sqrt{\frac{J+1}{2J+1}} g_{J-1}(kr) Y_{J,J-1,M}\left(\frac{r}{r}\right). \end{aligned}$$

Let us represent the electric field $\underline{E}$ and the magnetic field $\underline{H}$, as was done in (2.8), in the form of the real parts of complex vectors $2\mathcal{E}$ and $2\mathcal{H}$:

$$\begin{aligned} E(r) &= \mathcal{E}(r) + \mathcal{E}^*(r), \\ H(r) &= \mathcal{H}(r) + \mathcal{H}^*(r). \end{aligned}$$

The inverse Fourier transformation of expressions (4.26) and (4.27) gives:
for electric states

$$\begin{aligned} \mathcal{E}_{JM,M+1} &= \frac{i\sqrt{p^2 dp}}{4\pi^3} \left[\sqrt{\frac{J}{2J+1}} g_{J+1}(pr) Y_{J,J+1,M} + \right. \\ &\quad \left. + \sqrt{\frac{J+1}{2J+1}} g_{J-1}(pr) Y_{J,J-1,M} \right] e^{-ipr}, \\ \mathcal{H}_{JM,M+1} &= -\frac{\sqrt{p^2 dp}}{4\pi^3} g_J(pr) Y_{J,J,M} e^{-ipr}; \end{aligned} \quad (4.30)$$

for magnetic states

$$\left. \begin{aligned} \mathcal{E}_{pJM,0} &= \frac{i\sqrt{p^3 d\rho}}{4\pi^2} g_J(p\rho) Y_{JM} e^{-i\rho t}, \\ \mathcal{H}_{pJM,0} &= -\frac{\sqrt{p^3 d\rho}}{4\pi^2} \left[\sqrt{\frac{J}{2J+1}} g_{J+1} Y_{J,J+1,M} + \right. \\ &\quad \left. + \sqrt{\frac{J+1}{2J+1}} g_{J-1} Y_{J,J-1,M} \right] e^{-i\rho t}. \end{aligned} \right\} \quad (4.31)$$

We note that Equations (4.30) transform to (4.31) when we perform the substitution

$$\mathcal{E} \rightarrow -i\mathcal{H}; \quad \mathcal{H} \rightarrow i\mathcal{E}.$$

Expressions (4.30) and (4.31) are normalized so that there exists only one photon with an energy between p and $p + dp$. If the field is normalized according to (2.10), i.e., if we consider the case where one photon is found in a volume V which is chosen as a sphere of radius R , then the expressions for the fields will differ from (4.30) and (4.31) in the dp is replaced by π/R .

When normalized for a unit energy interval according to the condition

$$\int f_{pJM}^* f_{p'JM} dk = \delta(p - p')$$

the expressions for the fields will differ from (4.30) and (4.31) by the absence of the factor $\sqrt{dp}$.

In view of the transversality of the vector $\mathbf{Y}_{JM}^{(0)} = \mathbf{Y}_{JM}$, the magnetic field in electric states, and the electric field in magnetic states will be transverse, i.e.,

$$\mathbf{H}_{pJM,0} \cdot \mathbf{r} = \mathbf{E}_{pJM,0} \cdot \mathbf{r} = 0.$$

The electric field in electric states and the magnetic field in magnetic states, however, are expressed in terms of the vectors $\mathbf{Y}_{J,J \pm 1,M}$ and are therefore, not transverse. Expressions (4.13) and (4.17) can be used to express $\mathbf{Y}_{J,J \pm 1,M}$ in terms of the longitudinal vector $\mathbf{Y}_{JM}^{(-1)} = \mathbf{n} \cdot \mathbf{Y}_{JM}$ and the transverse one $\mathbf{Y}_{JM}^{(0)}$. We then obtain

$$\left. \begin{aligned} \mathcal{E}_{pJM,0} &= \frac{i\sqrt{p^3 d\rho}}{4\pi^2} \left\{ \frac{\sqrt{J(J+1)}}{2J+1} (g_{J-1} - g_{J+1}) Y_{JM}^{(-1)} + \right. \\ &\quad \left. + \left(\frac{1}{2J+1} g_{J+1} + \frac{J+1}{2J+1} g_{J-1} \right) Y_{JM}^{(0)} \right\} e^{-i\rho t} \\ \mathcal{H}_{pJM,0} &= i\mathcal{E}_{pJM,0}. \end{aligned} \right\} \quad (4.32)$$

The first term in (4.32) gives the radial component of the electric field. At large distances, when $pr \gg 1$ ("radiation zone"), the radial component vanishes since the asymptotic behaviors of g_{J-1} and g_{J+1} are identical (this is easily verified by using the asymptotic expressions for the Bessel functions) and the field becomes

$$\mathcal{E}_{pJM,0} = -i\mathcal{H}_{pJM,0} \approx \frac{i^{J+1}}{\sqrt{\pi}} \sqrt{p^3 d\rho} Y_{JM}^{(0)} \frac{\cos\left(pr - \pi - \frac{\pi}{2} J\right)}{r} e^{-i\rho t} \quad (pr \gg 1). \quad (4.33)$$

An arbitrary electromagnetic field can be expanded in a series of fields corresponding to angular momentum and parity eigenstates of the photon:

$$\mathcal{E} = \sum_{JM\lambda} C_{JM\lambda}^{pJM} \mathcal{E}_{pJM\lambda},$$

where the $C_{JM\lambda}^{pJM}$ are given by (4.23). In particular, the coefficients (4.24) define the expansion of plane polarized waves in terms of spherical waves:

$$\left. \begin{aligned} \mathcal{E}_{p\lambda}(r) &= \sum_{JM\lambda} \left(e_{\lambda} Y_{JM}^{(0)} \left(\frac{p}{r} \right) \right) \mathcal{E}_{pJM\lambda}(r) \sqrt{d\rho_p}, \\ \mathcal{H}_{p\lambda}(r) &= \sum_{JM\lambda} \left(e_{\lambda} Y_{JM}^{(0)*} \right) \mathcal{H}_{pJM\lambda}(r) \sqrt{d\rho_p}. \end{aligned} \right\} \quad (4.34)$$

§ 5. Potentials.

1. Transverse, Longitudinal, and Scalar Potentials.

In the future, in considering the interaction between photons and charges, we shall have need of expressions for the potentials of the electromagnetic field corresponding to a photon in some definite state. Therefore, we shall make use of the expressions obtained in the previous paragraph for the electromagnetic field of the photon to determine the vector $\mathbf{A}$ and scalar A_0 potentials of the fields $\mathbf{E}$ and $\mathbf{H}$ from the well-known expressions

$$\left. \begin{aligned} \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} - \nabla A_0, \\ \mathbf{H} &= \text{curl } \mathbf{A}. \end{aligned} \right\} \quad (5.1)$$

Let us perform a Fourier transformation, expressing $\mathbf{E}$ and $\mathbf{H}$ as in Equation (1.4), and the potentials $\mathbf{A}$ and A_0 in the form

$$\begin{aligned} \mathbf{A} &= \int \mathbf{A}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{k}, \\ A_0 &= \int A_0(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{k}. \end{aligned}$$

Then Equation (5.1) is transformed to

$$\left. \begin{aligned} E(k) &= -\dot{A}(k) - ikA_0(k), \\ H(k) &= i(kA(k)). \end{aligned} \right\} \quad (5.2)$$

Since $\underline{E}$ and $\underline{H}$ are transverse vectors, it is convenient to separate $\underline{A}$ into its transverse and longitudinal parts, namely

$$\underline{A} = \underline{B} + n\varphi \quad \left(n = \frac{k}{k} \right), \quad (5.3)$$

where

$$kB(k) = 0.$$

From (5.2), (5.3), and the transversality condition

$$nE = 0$$

of the field, it follows that $\underline{A}_0$ and φ are related by

$$ikA_0 + \dot{\varphi} = 0. \quad (5.4)$$

Eliminating the magnetic field with the aid of Equation (1.7), it is simple to obtain an expression relating the vectors $\underline{B}$ and $\underline{E}$ with $\underline{E}$ and $\underline{E}$:

$$\left. \begin{aligned} B &= \frac{1}{k^2} \dot{E}, \\ \dot{B} &= -E. \end{aligned} \right\} \quad (5.5)$$

From these expressions and (1.8), we can find expressions for $\underline{B}$ and $\dot{B}$ in terms of the photon wave function $f(k)$ in momentum space;

$$\left. \begin{aligned} B(k) &= \frac{-i}{4\pi^{1/2}\sqrt{k}} (f(k) - f^*(-k)), \\ \dot{B}(k) &= \frac{-i\sqrt{k}}{4\pi^{1/2}} (f(k) + f^*(-k)). \end{aligned} \right\} \quad (5.6)$$

The coefficients A_0 and φ are related only by Equation (5.4), remaining otherwise arbitrary. This is the expression of gauge invariance, according to which the fields $\underline{E}$ and $\underline{H}$ do not change if an arbitrary function $\frac{\partial \Lambda}{\partial t}$ is added to the scalar potential and at the same time $\nabla \Lambda$ is added to the vector potential. If the potentials are subjected to the subsidiary condition (Lorentz condition)

$$\text{div } \underline{A} + \frac{\partial A_0}{\partial t} = 0,$$

then $\underline{A}_0$ and φ will satisfy the relation

$$ik\varphi + \dot{A}_0 = 0. \quad (5.7)$$

Equations (5.4) and (5.7) for the scalars φ and $\underline{A}_0$ are analogous to the system of equations for the transverse vectors $\underline{E}$ and $\underline{H}$. We shall assume that $\underline{A}_0$ is real,

$$A_0(k) = A_0^*(-k).$$

Then, similarly as was done in (1.8), we can introduce the complex function $f_0(k)$ defined by

$$\left. \begin{aligned} A_0(k) &= \frac{1}{4\pi^{1/2}\sqrt{k}} (f_0(k) + f_0^*(-k)), \\ \dot{A}_0(k) &= \frac{-i\sqrt{k}}{4\pi^{1/2}} (f_0(k) - f_0^*(-k)) \end{aligned} \right\} \quad (5.8)$$

and satisfying the "Schrodinger equation"

$$i\frac{\partial f_0}{\partial t} = kf_0.$$

Then according to (5.3), (5.6) and (5.4), we obtain the following expression for $\underline{A}(k)$:

$$\underline{A}(k) = \frac{-i}{4\pi^{1/2}\sqrt{k}} (h(k) - h^*(-k)), \quad (5.9)$$

where

$$h(k) = f(k) + if_0(k). \quad (5.10)$$

The set of four quantities $\underline{h}_x, \underline{h}_y, \underline{h}_z, \underline{f}_0$, which determines the Fourier components of the potentials can serve as the photon wave function. Such a wave function, however, will satisfy no normalization condition. In fact, the expression for the expectation value of the photon energy now takes on the form

$$\overline{w} = \int k f^* f dk = \int k (h^* h - f_0^* f_0) dk,$$

and the quantity $\overline{w}$ is independent of $\underline{f}_0$ (compare Section 15).

Similarly, the expectation values of the momentum and angular momentum of the photon are expressed in terms of $\underline{h}$ and $\underline{f}_0$ by

$$\begin{aligned} \overline{p} &= \int k (h^* h - f_0^* f_0) dk, \\ \overline{M} &= \int (h^* M h - f_0^* L f_0) dk, \end{aligned}$$

where $\underline{M}$ is the angular momentum operator (3.6), and $\underline{L}$ is the orbital angular momentum operator (3.11).

2. Longitudinal and Scalar Components of the Photon Wave Function.

We obtain the potentials of the electromagnetic field corresponding to momentum and angular momentum eigenstates of the photon if we require that the wave function $(\underline{h}, \underline{f}_0)$ be an eigenfunction of the appropriate operators. Since we know the wave functions $\underline{f}$, we need only determine the function $\underline{f}_0$.

The eigenfunctions $\underline{f}_0$ of the momentum operator can be written in a form similar to (2.4), namely

$$f_{0p} = C \frac{1}{\sqrt{d\rho}} e^{-ip\rho} \delta_{kp}, \quad (5.11)$$

where C is an arbitrary constant. Thus, according to (5.10), the form of $\underline{h}$ for a definite momentum and polarization is

$$h_{pk} = \frac{ie^{-ip\rho}}{\sqrt{d\rho}} (e_k + iCn) \delta_{kp}. \quad (5.12)$$

The eigenfunctions of the angular momentum operator, which is identical with the orbital angular momentum operator for the scalar, are, according to (3.15) and (4.21),

$$f_{0jM} = \frac{iC}{\sqrt{d\rho}} Y_{jM} e^{-ip\rho} \delta_{kp}. \quad (5.13)$$

We note that the parity of the functions f_{0jM} is uniquely determined by the angular momentum, being equal to $(-1)^j$. This is the same as the parity of an electric state.

For a magnetic state, the scalar part of the wave function vanishes, and we obtain

$$\left. \begin{aligned} f_{0jM,0} &= 0, \\ h_{jM,0} &= f_{jM,0}. \end{aligned} \right\} \quad (5.14)$$

For an electric state

$$h_{pjM,+1} = \frac{i}{\sqrt{d\rho}} (Y_{jM}^+ + iCn Y_{jM}) e^{-ip\rho} \delta_{kp}. \quad (5.15)$$

Various choices of the arbitrary constant C give various forms for the function $\underline{h}_{jM,+1}$. For $C=0$, the vectors $\underline{h}$ and $\underline{f}$ are identical:

$$\left. \begin{aligned} h_{pjM,+1} &= f_{pjM,+1}, \\ f_{0pjM,+1} &= 0. \end{aligned} \right\} \quad (5.16)$$

For $C=1$, the vector $\underline{h}$ is proportional to one of the spherical functions $Y_{j-1,M}$.

$$\left. \begin{aligned} h_{pjM,+1} &= \frac{i}{\sqrt{d\rho}} \sqrt{\frac{2j+1}{j+1}} Y_{j,j-1,M} e^{-ip\rho} \delta_{kp}, \\ f_{0pjM,+1} &= \frac{i}{\sqrt{d\rho}} \frac{1}{\sqrt{j+1}} Y_{jM} e^{-ip\rho} \delta_{kp}. \end{aligned} \right\} \quad (5.17)$$

3. Plane and Spherical Wave Potentials

Returning from the wave functions $(\underline{h}, \underline{f}_0)$ to the Fourier components of the potentials (5.8), (5.9), and performing the inverse Fourier transformation, we obtain expressions for the plane wave potentials (momentum and polarization eigenstates of the photon) and spherical wave potentials (angular momentum and parity eigenstates).

As before, let us express the vectors $\underline{E}$ and $\underline{H}$ and the potentials $\underline{A}$, $\underline{A}_0$ in terms of the complex quantities $\underline{A}$ and $\underline{A}_0^*$ where

$$\begin{aligned} \underline{A}(r) &= \underline{A}(r) + \underline{A}^*(r), \\ \underline{A}_0(r) &= \underline{A}_0(r) + \underline{A}_0^*(r). \end{aligned}$$

It can be shown easily that the potentials $\underline{A}$ and $\underline{A}_0$ are of the following form:

for plane waves ¹⁾

$$\left. \begin{aligned} A_{pM} &= \frac{1}{4\pi^2} \sqrt{\frac{dp}{p}} \left(e_{\mu} + C \frac{p}{p} \right) e^{i(p\mathbf{r}-pt)}, \\ A_{0pM} &= \frac{1}{4\pi^2} \sqrt{\frac{dp}{p}} C e^{i(p\mathbf{r}-pt)}; \end{aligned} \right\} \quad (5.18)$$

for spherical waves of the electric type ²⁾

$$\left. \begin{aligned} A_{pJM, +1} &= \frac{\sqrt{p} dp}{4\pi^2} \left\{ \sqrt{\frac{J}{2J+1}} g_{J+1}(pr) Y_{J, J+1, M} \left(\frac{\mathbf{r}}{r} \right) + \right. \\ &\quad \left. + \sqrt{\frac{J+1}{2J+1}} g_{J-1} Y_{J, J-1, M} + \right. \\ &\quad \left. + C \left[\sqrt{\frac{J+1}{2J+1}} g_{J+1} Y_{J, J+1, M} - \sqrt{\frac{J}{2J+1}} g_{J-1} Y_{J, J-1, M} \right] \right\} e^{-ipt}, \\ A_{0pJM, +1} &= C \frac{\sqrt{p} dp}{4\pi^2} g_J Y_{JM} e^{-ipt}. \end{aligned} \right\} \quad (5.19)$$

When $C = 0$, ³⁾ the function A_{pJM} differs only by a factor $\frac{1}{Jp}$ from E_{pJM} . When

$C = -\sqrt{\frac{J}{J+1}}$ the expression for $A_{pJM, +1}$ contains only one term, which involves the function

E_{J-1} and is thus advantageous in many applications (see Sections 30 and 39):

$$\left. \begin{aligned} A_{pJM, +1} &= \frac{\sqrt{p} dp}{4\pi^2} \sqrt{\frac{2J+1}{J+1}} g_{J-1} Y_{J, J-1, M} e^{-ipt}, \\ A_{0pJM, +1} &= -\sqrt{\frac{J}{J+1}} \frac{\sqrt{p} dp}{4\pi^2} g_J Y_{JM} e^{-ipt}. \end{aligned} \right\} \quad (5.20)$$

For spherical waves of the magnetic type, we have for arbitrary C

$$\left. \begin{aligned} A &= \frac{1}{ip} \mathbf{E}; \\ A_{pJM, 0} &= \frac{\sqrt{p} dp}{4\pi^2} g_J Y_{JM}^{(0)}, \\ A_{0pJM, 0} &= 0. \end{aligned} \right\} \quad (5.21)$$

¹⁾ When normalizing for a volume V , the quantity dp should be replaced by $(2\pi)^3/V$ in (5.18).

²⁾ When normalizing for a sphere of radius R , the quantity dp in (5.19), (5.20), and (5.21) should be replaced by $(R/\pi)^{-1}$.

³⁾ This special case has been considered by Heitler [W. Heitler, Proc. Cambridge Phil. Soc. 32, 112 (1936)].

§ 6. The Two-photon System.

1. Two-photon wave function.

Up to this point we have been considering only a single photon and the electromagnetic field corresponding to it. The method of treating the wave function in momentum space makes it possible to consider also an arbitrary number of photons similarly as is done for a system of particles in quantum mechanics. In this section we shall undertake an investigation of the two-photon system.

The arguments of the two-photon wave function are the momenta $\mathbf{k}_1, \mathbf{k}_2$ and spin variables α_1, α_2 of both particles. We shall write the wave function

$$f(\mathbf{k}_1, \alpha_1; \mathbf{k}_2, \alpha_2).$$

The square of its absolute value determines the probability of finding one photon with a momentum $\mathbf{k}_1$ and polarization in the direction given by $\alpha_1 (\alpha_1 = \mathbf{k}_1, \mathbf{y}, \mathbf{z})$, and the other with a momentum $\mathbf{k}_2$ and polarization given by $\alpha_2 (\alpha_2 = \mathbf{k}_2, \mathbf{y}, \mathbf{z})$. Instead of representing the wave function as a scalar in the space of all the variables $\mathbf{k}_1, \mathbf{k}_2, \alpha_1, \alpha_2$, we can consider it a tensor function in the space of momenta $\mathbf{k}_1, \mathbf{k}_2$, so that:

$$f(\mathbf{k}_1, \alpha_1; \mathbf{k}_2, \alpha_2) = f_{\alpha_1 \alpha_2}(\mathbf{k}_1, \mathbf{k}_2). \quad (6.1)$$

Just as we treated the one-photon wave function (see Section 3) either as a scalar $f(\mathbf{k}, \alpha)$ or as a vector $f_\alpha(\mathbf{k})$.

In the first approximation the photons can be considered noninteracting particles.⁴⁾ Then the energy of a two-photon system is the sum of the energies of each photon, and their wave function satisfies the Schrodinger equation

$$i \frac{\partial f}{\partial t} = (k_1 + k_2) f. \quad (6.2)$$

In addition to this equation, the wave function must satisfy two other conditions. The first of these is the transversality condition for each photon; it can be formulated by the relations

$$\left. \begin{aligned} \sum_{\alpha_1} (\mathbf{k}_1)_{\alpha_1} f_{\alpha_1 \alpha_2} &= 0, \\ \sum_{\alpha_2} (\mathbf{k}_2)_{\alpha_2} f_{\alpha_1 \alpha_2} &= 0. \end{aligned} \right\} \quad (6.3)$$

If conditions (6.3) are satisfied, the probability vanishes for polarization of the first photon along $\mathbf{k}_1$ or of the second along $\mathbf{k}_2$.

The second condition is a symmetry condition which follows from the identity of the photons. Photons satisfy Bose-Einstein statistics and their wave function should therefore be symmetric with respect to exchange of the particles. Thus,

⁴⁾ We shall see (see Section 47) that in view of the interaction of photons with electrons, there exists also a weak interaction of photons with each other.

$$f(k_1, \alpha_1; k_2, \alpha_2) = f(k_2, \alpha_2; k_1, \alpha_1), \quad (6.4)$$

or in tensor form

$$f_{\alpha_1 \alpha_2}(k_1, k_2) = f_{\alpha_2 \alpha_1}(k_2, k_1).$$

Instead of the variables $\underline{k}_1, \underline{k}_2$ we can introduce the total momentum of the system (the momentum of the center of mass)

$$\underline{K} = \underline{k}_1 + \underline{k}_2$$

and the relative momentum $\underline{k}$, where

$$\underline{k} = \frac{\underline{k}_1 - \underline{k}_2}{2}.$$

In these variables the wave function of the system can be written

$$f(k_1, \alpha_1; k_2, \alpha_2) = \varphi(\underline{K}) f(\underline{k}; \alpha_1, \alpha_2).$$

The function $f(\underline{k}; \alpha_1, \alpha_2) = f_{\alpha_1 \alpha_2}(\underline{k})$, in view of the transversality condition (6.3) and symmetry condition (6.4), satisfies the relations

$$\left. \begin{aligned} \sum_{\alpha_1} f_{\alpha_1 \alpha_2} k_{\alpha_1} &= 0, \\ \sum_{\alpha_2} f_{\alpha_1 \alpha_2} k_{\alpha_2} &= 0; \end{aligned} \right\} \quad (6.5)$$

$$f(\underline{k}; \alpha_1, \alpha_2) = f(-\underline{k}; \alpha_2, \alpha_1). \quad (6.6)$$

When considering a state with a definite total momentum $\underline{K}$, it is always possible (except when $\underline{k}_1$ and $\underline{k}_2$ are parallel) to transform to a coordinate system in which

$$\underline{K} = 0; \quad \underline{k}_1 = \underline{k}; \quad \underline{k}_2 = -\underline{k}.$$

The function $f(\underline{k}; \alpha_1, \alpha_2)$ is a two-photon wave function in this coordinate system. Two photons with zero total momentum are experimentally observed in the decay of electrically neutral systems at rest (neutral meson, positronium).

2. Even and Odd States.

Let us attempt to find the wave functions of states with definite angular momentum and parity for the two-photon system¹⁾ whose total momentum $\underline{K} = 0$. The angular momentum operator $\underline{M}$ of the system is the sum of the orbital and spin momenta:

$$\underline{M} = \underline{L} + \underline{s},$$

where

$$\underline{L} = -i[\underline{k}, \nabla_{\underline{k}}]$$

is the orbital angular momentum of the relative motion, and

$$\underline{s} = \underline{s}_1 + \underline{s}_2$$

is the sum of the spin operators of both photons.

The eigenfunctions of the operators $\underline{L}^2$ and $\underline{L}_z$ are, as in the case of a single photon, the spherical functions

$$\Phi_{lm} = a(l) Y_{lm}(\pi).$$

The eigenvalues of the square of the spin angular momentum are

$$s^2 = s(s+1),$$

where, according to the rules for composition of angular momenta, the number s may take on the values

$$s = 0, 1, 2.$$

Since for a given value of $\underline{s}$, the eigenvalue of the $\underline{z}$ component of the spin

$$s_z = \mu$$

¹⁾ L. Landau, Proc. Acad. Sci. USSR 60, 207 (1948).

may take on the values

$$\mu = -s, -s+1, \dots, s,$$

and there exist only nine different spins states of a two-photon system.

The spin wave functions

$$\chi_{\mu_1 \mu_2}(\alpha_1, \alpha_2)$$

of the system are bilinear combinations of the spin wave functions of each of the photons:

$$\chi_{\mu_1 \mu_2}(\alpha_1, \alpha_2) \quad (\mu_1, \mu_2 = 0, \pm 1).$$

Let us construct the following normalized combinations which have definite symmetries with respect to exchange of the spin variables α_1, α_2 :

$$\left. \begin{aligned} &\chi_{\mu_1 \mu_2}(\alpha_1) \chi_{\mu_1 \mu_2}(\alpha_2) \quad (\mu_1 = \mu_2; 3 \text{ functions}), \\ &\frac{1}{\sqrt{2}} \{ \chi_{\mu_1 \mu_2}(\alpha_1) \chi_{\mu_2 \mu_1}(\alpha_2) + \chi_{\mu_2 \mu_1}(\alpha_1) \chi_{\mu_1 \mu_2}(\alpha_2) \} \quad (\mu_1 \neq \mu_2; 3 \text{ functions}), \\ &\frac{1}{\sqrt{2}} \{ \chi_{\mu_1 \mu_2}(\alpha_1) \chi_{\mu_2 \mu_1}(\alpha_2) - \chi_{\mu_2 \mu_1}(\alpha_1) \chi_{\mu_1 \mu_2}(\alpha_2) \} \quad (\mu_1 \neq \mu_2; 3 \text{ functions}). \end{aligned} \right\} \quad (6.7)$$

The first six functions in (6.7) are symmetric, and the last three are antisymmetric. Since the states with different quantum numbers μ and the same quantum number $\underline{1}$ should have the same symmetry, we can identify the three antisymmetric spin functions with the functions $\chi_{\underline{1}\mu}(\mu = \mu_1 + \mu_2)$ for $\underline{1}=1$.

The six symmetric functions (6.7) refer to $s=0, 2$.

Of these, the four functions for which $\mu_1 + \mu_2 \neq 0$ can be identified with the wave function $\chi_{\underline{1}\mu}$ for $\underline{1}=2$, $\mu = \mu_1 + \mu_2$. The other two functions

$$\chi_0(\alpha_1) \chi_0(\alpha_2)$$

and

$$\frac{1}{\sqrt{2}} \{ \chi_1(\alpha_1) \chi_{-1}(\alpha_2) + \chi_{-1}(\alpha_1) \chi_1(\alpha_2) \}$$

are linear combinations of the wave functions $\chi_{\underline{2}\mu}$ and $\chi_{\underline{0}\mu}$.

We note that the separation of the spin states into symmetric and antisymmetric has actual significance. As for the classification of states in terms of spin values, it has no profound physical meaning since, as will be seen later, the transversality condition requires a superposition of states with $\underline{1}=0$ and $\underline{1}=2$.

The wave function of the two-photon system which corresponds to definite values of the angular momentum $\underline{1}$ and its projection $\underline{M}$, namely

$$f_{\underline{1}\underline{M}}(\underline{k}; \alpha_1, \alpha_2)$$

is, according to the rules for composition of angular momenta, a linear combination of products of orbital and spin functions

$$Y_{\underline{1}\underline{M}}(\underline{n}) \chi_{\mu_1 \mu_2}(\alpha_1, \alpha_2).$$

Eigenstates of the angular momentum can be further classified in terms of their parity. The effect of the inversion operator I on the tensor function is determined in the following way:¹⁾

$$I f_{\alpha_1 \alpha_2}(\underline{k}) = f_{\alpha_1 \alpha_2}(-\underline{k}),$$

that is,

$$I f(\underline{k}; \alpha_1, \alpha_2) = f(-\underline{k}; \alpha_1, \alpha_2).$$

Thus, the inversion operator acts only on the orbital part of the wave function. The parity of the state is determined by $(-1)^{\underline{1}}$, so that

$$Y_{\underline{1}\underline{M}}(-\underline{n}) = (-1)^{\underline{1}} Y_{\underline{1}\underline{M}}(\underline{n}).$$

On the other hand, replacing $\underline{k}$ by $-\underline{k}$ is equivalent to interchanging the momenta of the two photons. According to the symmetry condition (6.8) of the wave function, we have

$$Y_{\underline{1}\underline{M}}(\underline{n}) \chi_{\mu_1 \mu_2}(\alpha_1, \alpha_2) = Y_{\underline{1}\underline{M}}(-\underline{n}) \chi_{\mu_2 \mu_1}(\alpha_2, \alpha_1).$$

We see, therefore, that there is a relation between the parity of a state and the symmetry properties of the spin function:

$$\chi_{\mu_1 \mu_2}(\alpha_1, \alpha_2) = (-1)^{\underline{1}} \chi_{\mu_2 \mu_1}(\alpha_2, \alpha_1), \quad (6.8)$$

¹⁾ Since under inversion all the directions are reflected, the components of a tensor of odd rank change sign (see Section 4), and those of an even rank do not.

that is, even states ($I = 2n$, where n is any integer) are symmetric with respect to the spin variables, and odd states ($I = 2n+1$) are antisymmetric. In other words, the tensor $f_{\alpha_1\alpha_2}(\mathbf{k})$ is a symmetric tensor in even states, and is antisymmetric in odd states. In the former case it has six components, and in the second case, three, and these correspond to the number of symmetric and antisymmetric functions (6.7).

3. Classification of States With a Given Angular Momentum.

Let us start with an analysis of the odd states. An antisymmetric tensor of second rank, as is well known, can be written in the form

$$f_{\alpha_1\alpha_2} = \sum_i \epsilon_{\alpha_1\alpha_2 i} f_i, \quad (6.9)$$

where $\epsilon_{\alpha_1\alpha_2 i}$ is the unit antisymmetric tensor, and f_i is a vector. Therefore, odd states can be described by a vector wave function $\mathbf{f}(\mathbf{k})$. According to the results of Section 4, there exist three linearly independent vector functions $f_{\mathbf{M}}^{(\lambda)}$, corresponding to given values of the quantum numbers $\mathbf{j}$ and $\mathbf{M}$, such that two of these are transverse ($\lambda = 0, 1$), and one is longitudinal ($\lambda = -1$). The transversality condition (6.5) when applied to the antisymmetric tensor (6.9) becomes

$$[\mathbf{k}|\mathbf{f}] = 0.$$

Thus, the vector $f_{\mathbf{M}}$ should be longitudinal, and for given $\mathbf{j}$ and $\mathbf{M}$ there therefore exists only one odd state with $\lambda = -1$:

$$f_{\mathbf{M}} = a Y_{\mathbf{M}}^{(-1)}.$$

According to (4.14) and (4.13), $Y_{\mathbf{M}}^{(-1)}$ contains the values $l = j \pm 1$. Since l is odd in the case we are considering, j must be even:

$$j = 2n; \quad l = 2n \pm 1. \quad (6.10)$$

We note that since the tensor wave function $f_{\alpha_1\alpha_2}$ of the two-photon system is bilinear in the components of the photon vector functions $\mathbf{f}_1$ and $\mathbf{f}_2$, the antisymmetry of $f_{\alpha_1\alpha_2}$ means that the vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ do not have components along the same axis. Since the component f_{α} determines the probability for polarization along the axis α , this means that two photons in an odd state are polarized perpendicular to each other. We note that a system decaying into two photons, for instance a neutral meson or positronium (see Section 33) in the singlet ground state, is odd, and the above results are applicable to these decays.

Let us go on to a consideration of even states. Since they correspond to the values $j = 0, 2$, then according to the rules for the composition of angular momenta with given $\mathbf{j}$ and $\mathbf{M}$, there are six different wave functions corresponding to values of l given by

$$\begin{aligned} l &= j \pm 1; \quad j \pm 2; \quad (s=2), \\ l &= j \quad (s=0). \end{aligned}$$

Since in this case l can take on only even values, the number of wave functions is given by

$$\begin{aligned} 4 &\text{ for } j = 2n, \\ 2 &\text{ for } j = 2n+1. \end{aligned}$$

In addition, we must take account of the transversality condition (6.3). To do this we shall make use of the possibility of constructing functions $(f_{\mathbf{M}}^{(\lambda)})_{\alpha_1\alpha_2}$ ("spherical tensors") in the form of bilinear combinations of spherical vectors:

$$(f_{\mathbf{M}}^{(\lambda)})_{\alpha_1\alpha_2} \sim (Y_{\mathbf{j}_1}^{(\lambda_1)})_{\alpha_1} (Y_{\mathbf{j}_2}^{(\lambda_2)})_{\alpha_2}. \quad (6.11)$$

Here $\mathbf{M}_1 + \mathbf{M}_2 = \mathbf{M}$ and the values $\mathbf{j}_1$ and $\mathbf{j}_2$ should be such that the value $\mathbf{j}$ can be obtained from them according to the rules for composition of angular momentum. Corresponding to the three values of λ_1 and λ_2 one can obviously construct six different symmetric bilinear combinations, which correspond to the six tensor wave functions $(f_{\mathbf{M}}^{(\lambda)})_{\alpha_1\alpha_2}$. Of these six combinations, three do not contain a longitudinal vector and satisfy the transversality condition. The remaining three functions ("longitudinal") cannot correspond to actual photons.

In order to find the parity of the functions which satisfy the transversality condition, let us construct an explicit expression for the longitudinal spherical tensors. We shall choose $j_1 = 0$ in (6.11), without loss of generality. Then

$$j_2 = j; \quad M_2 = M; \quad M_1 = 0,$$

and since [see (4.14)]:

$$Y_{00}^{(-1)} = n Y_{00},$$

the longitudinal spherical tensors will be of the following form:

$$(f_{\mathbf{M}}^{(\lambda)})_{\alpha_1\alpha_2} = n_{\alpha_1} (Y_{\mathbf{j}}^{(\lambda)})_{\alpha_2} + n_{\alpha_2} (Y_{\mathbf{j}}^{(\lambda)})_{\alpha_1}, \quad \lambda = 0, \pm 1. \quad (6.12)$$

When $\lambda = \pm 1$, the function $Y_{\mathbf{M}}^{(\lambda)}$ contains [see (4.13), (4.17)] spherical functions with l values given by $l = j \pm 1$; when $\lambda = 0$, these are given by $l = j$. Since (6.12) contains the vector $\mathbf{n}$, and $\mathbf{l} = -\mathbf{n}$, we obtain

$$\begin{aligned} l f_{\mathbf{M}}^{(0)} &= (-1)^j f_{\mathbf{M}}^{(0)} & \text{for } \lambda = \pm 1, \\ l f_{\mathbf{M}}^{(\pm 1)} &= (-1)^{j \pm 1} f_{\mathbf{M}}^{(\pm 1)} & \text{for } \lambda = 0. \end{aligned}$$

This means that the number of even longitudinal states is

$$\begin{aligned} 2 & \text{ for } j=2n, \\ 1 & \text{ for } j=2n+1. \end{aligned}$$

Subtracting this number from the total number of even states, we obtain the number of even states which satisfy the transversality condition. There are, then,

$$\begin{aligned} 2 & \text{ for } j=2n, \\ 1 & \text{ for } j=2n+1. \end{aligned}$$

The cases $j=0$ and $j=1$ require separate consideration. When $j=0$, the rule for composition of angular momenta allows only two states (both even):

$$l=0; s=0 \text{ and } l=2; s=2.$$

On the other hand, there exists only one spherical vector with $j=0$ (longitudinal). Therefore, only the value $\lambda=0$ is possible in Equation (6.12). This means that of the two tensors $(f_{\mu\nu})$ $\alpha_\mu \alpha_\nu$, one is longitudinal and the other is therefore transverse. Thus, there exists only one even state of the system with zero angular momentum. In this case, as is not difficult to prove, the wave function which satisfies the symmetry and transversality conditions is

$$(f_{00})_{\mu\nu} = a(k) (n_\mu n_\nu - \delta_{\mu\nu}), \quad (6.13)$$

One may conclude from (6.13) that in this case the photons are polarized parallel to each other. Indeed, let the z axis be directed along k ; then $f_{xy}=0$;

When $j=1$, only one even state is possible; $l=2, s=2$. But an even state is already contained among the functions of (6.12) (namely $(f_{\mu\nu}) \alpha_\mu \alpha_\nu$), which do not satisfy the transversality condition. This means that there exist no even states with $j=1$. According to (6.10), there also exist no odd states with $j=1$. Thus, a two-photon system can never be in a state with unit angular momentum, and a system with unit angular momentum cannot decay into two photons.

In conclusion we present a summary of the classification of two-photon states.

Angular Momentum	No. of even states	No. of odd states
0	1	1
1	2	1
2n	1	—
2n+1	—	—

CHAPTER II

RELATIVISTIC QUANTUM MECHANICS OF THE ELECTRON

§ 7. The Dirac Equation.

1. Spinors and Pauli Matrices.

In nonrelativistic quantum mechanics the motion of the electron is described by a two-component wave function, namely a spinor $\varphi(\mathbf{r}, t)$. The components of the spinor are usually denoted by φ^λ , where the index λ takes on the values $\lambda = \pm 1/2$. The spinor is written in the form of a column vector

$$\varphi = \begin{pmatrix} \varphi^{1/2} \\ \varphi^{-1/2} \end{pmatrix}. \quad (7.1)$$

Under a rotation of the coordinate axes, the spinor components transform according to¹⁾

$$\varphi = \left(1 + \frac{i}{2} \delta \sigma\right) \varphi', \quad (7.2)$$

where φ' denotes the spinor φ (at the same point $\mathbf{r}$ in space), except that its components are given in terms of a coordinate system rotated relative to the original one by an infinitesimal angle δ about the axis directed along the vector δ , and σ is the set of Pauli matrices²⁾

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.3)$$

According to the general relation between the infinitesimal rotation operator and the angular momentum, the matrix $\frac{1}{2} \sigma$ is the intrinsic angular momentum operator of the electron, or the spin. The transformation law (7.2) is a result of the fundamental physical fact that the electron spin is $\frac{1}{2}$.

¹⁾ See, for instance, L. Landau and E. Lifshits, Quantum Mechanics, Section 54.

²⁾ The expressions for the σ_i ($i = x, y, z$) are independent of the coordinate system. Therefore, σ is not a vector. Nevertheless, vector notation, which is extremely convenient, leads to no confusion.

The matrices (7.3) satisfy the commutation rules of the components of the angular momentum, namely

$$\sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z \quad (7.4)$$

(and similar expressions obtained by cyclic permutation of the indices). In addition, they satisfy the relation

$$\sigma_i \sigma_k + \sigma_k \sigma_i = 2\delta_{ik} \quad (7.5)$$

It follows from (7.4) and (7.5) that

$$\sigma_i \sigma_k = \delta_{ik} + i\epsilon_{ikl} \sigma_l \quad (7.6)$$

where ϵ_{ijk} is the unit antisymmetric tensor.

As do other types of quantities (for instance tensors), spinors differ in the way they transform under reflections. A spinor φ which transforms according to

$$\varphi = \varphi' \quad (7.7)$$

where φ' is the same spinor, but with its components given in the reflected coordinate system, shall be called a polar spinor (or simply a spinor). A spinor χ , on the other hand, which under reflection transforms according to

$$\chi = -\chi' \quad (7.8)$$

shall be called a pseudospinor. In either case, double reflection leads to the identity transformation.

We note that as opposed to tensors, spinors can have transformation rules under reflection¹⁾ other than (7.7) or (7.8). This is because according to (7.2), under a rotation of 2π , the spinor components transform according to²⁾

$$\varphi^{\lambda} = e^{2\pi i} \varphi^{\lambda'} = -\varphi^{\lambda'}$$

Thus, if we define the double reflection as a rotation through 2π , rather than the identity transformation, under reflection the spinor transforms according to

¹⁾ See W. Pauli, Relativistic Theory of Elementary Particles (Foreign Lit. Press, 1947); G. Zharikov, J. Exptl.-Theoret. Phys., 20, 492 (1950); Yang and Tiomno, Phys. Rev., 79, 495 (1951).

²⁾ Indeed, it can be shown from (7.2) and (7.3) that under rotation through a finite angle δ about the z axis, the transformation is given by

$$\varphi^{1/2} = e^{i\frac{\delta}{2}} \varphi^{1/2}, \quad \varphi^{-1/2} = e^{-i\frac{\delta}{2}} \varphi^{-1/2}$$

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi^1 \\ \varphi^2 \end{pmatrix} \quad (7.8')$$

We shall henceforth use the first definition of the double reflection, and thus, use the transformations given by (7.7) and (7.8').

2. Dirac Equation.

The relativistic equations which perform the same function for the electron as Maxwell's equations do for the photon were found by Dirac.¹⁾ These are the following system of first order homogeneous differential equations for the two spinor functions φ and χ ²⁾

$$\left. \begin{aligned} i \frac{\partial \varphi}{\partial t} &= m\varphi + \sigma p \chi, \\ i \frac{\partial \chi}{\partial t} &= -m\chi + \sigma p \varphi. \end{aligned} \right\} \quad (7.9)$$

Here $p = -i\hbar \nabla$ is the momentum operator, and m is the electron mass.

The set of two spinors φ and χ can be written in the form of a single four-component quantity

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} \varphi^{1/2} \\ \varphi^{-1/2} \\ \chi^{1/2} \\ \chi^{-1/2} \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix},$$

which we shall call a bispinor.

In order to write (7.9) in the form of a single equation for ψ , we introduce the four-dimensional matrices

¹⁾ P. Dirac, Quantum Mechanics (United Sci. Tech. Press, 1937).

²⁾ We note that Maxwell's equations can also be written in a form similar to (7.9), namely

$$\frac{\partial E}{\partial t} = (sp) H, \quad \frac{\partial H}{\partial t} = -(sp) E,$$

where s is the photon spin operator (3.3) (the photon mass is zero).

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.10)$$

(The symbols 1 in the expression for β represent two-dimensional unit matrices.) Using this notation, we can write (7.9) in the form¹⁾

$$i \frac{\partial \psi}{\partial t} = (\alpha p + \beta m) \psi. \quad (7.11)$$

Equation (7.11) is in the form of a Schroedinger equation,

$$i \frac{\partial \psi}{\partial t} = H \psi,$$

in which the Hamiltonian is the expression

$$H = \alpha p + \beta m. \quad (7.12)$$

Let us point out the properties of the matrices α and β . From the definitions (7.10) and from (7.5), we obtain

$$\left. \begin{aligned} \alpha_i \alpha_k + \alpha_k \alpha_i &= 2\delta_{ik}, \\ \beta^2 &= 1, \\ \alpha\beta + \beta\alpha &= 0. \end{aligned} \right\} \quad (7.13)$$

In the future it will be convenient also to introduce the following additional four-dimension matrices

¹⁾ In this equation we use ordinary matrix multiplication, so that

$$\begin{aligned} (\alpha\psi)_i &= \sum_{k=1}^4 \alpha_{ik} \psi_k, \\ (\psi^* \alpha)_i &= \sum_{k=1}^4 \psi_k^* \alpha_{ki}. \end{aligned}$$

Similarly,

$$\psi^* \psi = \sum_{i=1}^4 \psi_i^* \psi_i.$$

$$\Sigma = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}, \quad \rho = \begin{pmatrix} 0 & \sigma \\ \sigma & 1 \end{pmatrix}. \quad (7.14)$$

Here, as in (7.10), the number 1 denotes the two-dimensional unit matrix.²⁾ The matrices Σ satisfy the same equations (7.4), (7.5), and (7.6) as do the σ_i . In addition, as can be seen from the appropriate definitions, the following relations hold:

$$\left. \begin{aligned} \rho \Sigma - \Sigma \rho &= 0, \\ \beta \Sigma - \Sigma \beta &= 0, \\ \rho \beta - \beta \rho &= 0, \\ \rho^2 &= 1. \end{aligned} \right\} \quad (7.15)$$

The relations between the matrices α_i and Σ_k are easy to obtain if we bear in mind that

$$\alpha = \rho \Sigma. \quad (7.16)$$

Thus, replacing σ by Σ in (7.4), (7.5), and (7.6), we arrive at

$$\left. \begin{aligned} \sigma_x \sigma_y - \sigma_y \sigma_x &= 2i \Sigma_z, \\ \sigma_i \Sigma_k &= \Sigma_i \sigma_k = \delta_{ik} \rho + i \epsilon_{ikl} \sigma_l, \\ [\Sigma \Sigma] &= 2i \alpha. \end{aligned} \right\} \quad (7.17)$$

3. On the Necessity for Four-Component Wave Functions.

Let us make some remarks that will clarify the structure of the Dirac equation. The transition from the spinor φ to the bispinor ψ is analogous to transition from the three-dimensional vector $\underline{a}$ to the four-dimensional vector $\underline{a} = (\underline{a}; a_0)$ (a_0 is the time component of the four-dimensional vector). Just as $\underline{a}$ and $\underline{a}_0$ transform independently of each other under space rotation;

$$\left. \begin{aligned} \underline{a} &= \underline{a}' - [\delta \underline{a}'], \\ a_0 &= a'_0 \end{aligned} \right\} \quad (7.18)$$

²⁾ We note that when the number 1 is replaced by these two-dimensional unit matrices, σ_x is transformed into ρ and σ_x into β .

(δ is the infinitesimal rotation vector, as in (7.2)), so the two spinors φ and χ which combine to make up the bispinor ψ transform independently according to Equation (7.2) under rotation. Under a general Lorentz transformation, however, ψ and δ no longer transform independently. If δ_0 is an infinitesimal vector giving the velocity of the new coordinate system relative to the original one, then

$$\left. \begin{aligned} \alpha &= \alpha' + \frac{1}{2} \delta_0 \alpha', \\ a_0 &= a'_0 + \delta_0 \alpha'. \end{aligned} \right\} \quad (7.19)$$

It is easy to see that the spinors and φ and χ cannot transform independently under general Lorentz transformations. In fact, the four matrices σ_x , σ_y , σ_z and 1 are a complete linearly independent set of two-dimensional matrices. Therefore, a transformation linear in δ_0 which reduces to the identity when $\delta_0 = 0$, can only be of the form

$$\varphi = \left(1 + \frac{\nu}{2} \delta_0 \sigma\right) \varphi', \quad (7.20)$$

where ν is a constant. But such a transformation equation cannot be valid, since it is not invariant with respect to reflection. To show this, let us compare it with (7.2). In that equation the infinitesimal rotation vector δ is a pseudovector, whereas the infinitesimal velocity vector δ_0 in (7.20) is a polar vector. Therefore, if (7.2) is to be valid in two coordinate systems which are reflections of each other, then the second term in (7.20) has different signs in these two systems. At the same time the set of two spinors φ and χ , one of which is a pseudospinor, can transform under Lorentz transformation according to

$$\left. \begin{aligned} \varphi &= \varphi' + \frac{\nu}{2} (\delta_0 \sigma) \chi', \\ \chi &= \chi' + \frac{\nu}{2} (\delta_0 \sigma) \varphi'. \end{aligned} \right\} \quad (7.21)$$

We thus see that the electron wave function must consist of two spinors (one bispinor).

4. Invariance of the Dirac Equation.

Let us now go on to a proof that the Dirac equation is invariant with respect to Lorentz transformations. Let us consider inversions, rotations, and actual Lorentz transformations in turn.

1) Inversions. If the bispinor ψ consists of a spinor φ and a pseudospinor χ , then (7.7) and (7.8) can be written

$$\psi = \beta \psi', \quad (7.22)$$

where β is the matrix given by (7.10). Introducing the notation

$$i \frac{\partial}{\partial t} = p_0,$$

we write the Dirac equation (7.11) in the form

$$(p_0 - \alpha p - \beta m) \psi = 0. \quad (7.23)$$

Let us now perform an inversion of the space axes. Then p_0 and m , since they are scalars, do not change, so that $p_0 = p'_0$; the numerical matrices β and α are independent of the coordinate system and also remain invariant. The components of the vector p , however, change sign under inversion; therefore,

$$\alpha p = -\alpha p'.$$

Thus, Equation (7.23) is transformed to

$$(p'_0 + \alpha p' - \beta m) \psi' = 0.$$

Let us multiply this equation on the left by β . With the aid of (7.13) we obtain

$$(p'_0 - \alpha p' - \beta m) \psi' = 0,$$

which is the same as Equation (7.23), but in the transformed coordinate system.

2) Rotations. Both spinors which go to make up ψ transform according to (7.2). From the definition (7.14) we can write the transformation law in the form

$$\psi = \left(1 + \frac{i}{2} \delta \Sigma\right) \psi'. \quad (7.24)$$

It follows from (7.24) that the matrix $\frac{1}{2} \Sigma$ is the intrinsic angular momentum (spin) operator of the electron. Let us insert (7.24) into the Dirac equation (7.23). Then since p is a vector, it transforms according to (7.18), and we obtain

$$[p'_0 - \alpha(p' - [\delta p'])] - \beta m \left(1 + \frac{i}{2} \delta \Sigma\right) \psi' = 0.$$

Let us multiply this equation on the left by $(1 - \frac{i}{2} \delta \Sigma)$ and neglect all terms quadratic in δ . We then obtain

$$\begin{aligned} (p'_0 - \alpha p' - \beta m) \psi' = \\ = -\frac{i}{2} m \delta (\Sigma \beta - \beta \Sigma) - \frac{i}{2} \{[\delta \Sigma] (\alpha p') - (\alpha p') [\delta \Sigma]\} \psi' - [\delta p'] \alpha \psi'. \end{aligned}$$

The first term on the right vanishes according to (7.15). The second term can be rewritten with the aid of (7.17), so that

$$-\frac{i}{2} \{ (\partial \Sigma)(\alpha p') - (\alpha p')(\partial \Sigma) \} = (\partial p') \alpha.$$

This cancels with the third term, and we arrive again at Equation (7.23).

3) Lorentz transformation. The transformation (7.21) with $\nu = 1$ can be written

$$\psi = \left(1 + \frac{1}{2} \partial_0 \alpha\right) \psi'. \quad (7.25)$$

Inserting (7.25) into (7.23), and using Equation (7.19) to transform p and p_0 , we arrive at

$$[p'_0 + \partial_0 p' - \alpha(p' + \partial_0 p'_0) - \beta m] \left(1 + \frac{1}{2} \partial_0 \alpha\right) \psi' = 0.$$

Multiplying this equation on the left by $(1 + \frac{1}{2} \partial_0 \alpha)$, we obtain, to terms of first order in $\partial_0 \alpha$,

$$\begin{aligned} (p'_0 - \alpha p' - \beta m) \psi' &= \\ &= \frac{1}{2} m \partial_0 (\alpha^2 + \beta \alpha) \psi' + \frac{1}{2} \{ (\alpha p') (\partial_0 \alpha) + (\partial_0 \alpha) (\alpha p') \} \psi' - (\partial_0 p') \psi'. \end{aligned}$$

According to (7.13) the first term on the right vanishes, and the second term (that in the braces) cancels with the last one, so that we again arrive at Equation (7.23).

5. The γ Matrices. Continuity Equation.

The Dirac equation can be written in a more elegant form if we introduce the four matrices γ_j ($j = 1, 2, 3, 4$), which are defined by

$$\left. \begin{aligned} \gamma_j &= -i\beta \alpha_j \quad (j = 1, 2, 3), \\ \gamma_4 &= \beta. \end{aligned} \right\} \quad (7.26)$$

We shall denote the set of three matrices $\gamma_1, \gamma_2, \gamma_3$ by the symbol γ . The matrices γ_j satisfy the following relations:

$$\gamma_i \gamma_k + \gamma_k \gamma_i = 2\delta_{ik} \quad (i, k = 1, 2, 3, 4). \quad (7.27)$$

Just as all the matrices previously introduced, the γ_j are Hermitian, so that

$$\gamma_i^\dagger = \gamma_i, \quad \gamma_4^\dagger = \gamma_4 \quad (7.28)$$

($\gamma_j^\dagger$ is the Hermitian conjugate of the matrix, and $\tilde{\gamma}_j$ is the transposed matrix).

Let us multiply the Dirac equation (7.23) on the left by $\tilde{\gamma}_j$. Using the definitions (7.26), we obtain

$$(\tilde{\gamma}_j p_0 + \gamma_j p - im) \psi = 0.$$

If we introduce the abbreviated notation

$$\sum_j \gamma_j p_j = \gamma p + \gamma_4 p_4 = \gamma p, \quad (7.29)$$

for the "scalar product", where

$$p_4 = ip_0,$$

then the Dirac equation takes on the form

$$(i\gamma p + m) \psi = 0. \quad (7.30)$$

The operator γp in (7.30) is not self-conjugate. Therefore, the complex conjugate wave function ψ^* satisfies an equation which is not the same as (7.30). If we introduce the function $\bar{\psi}$, defined by

$$\bar{\psi} = \psi^* \beta, \quad \psi^* = \bar{\psi} \beta, \quad (7.31)$$

then $\bar{\psi}$ will satisfy an equation that can be written in the form

$$\bar{\psi} (-i\gamma p + m) = 0. \quad (7.32)$$

In (7.32) the differential operators p_j are assumed to act on the function $\bar{\psi}$ to their left.

Let us prove the validity of Equation (7.32). Since

$$p_4^* = -p_4$$

and

$$i\hbar \frac{\partial}{\partial t} \psi = H \psi,$$

the expression which is the complex conjugate of (7.30) will be

$$(-i\hbar \frac{\partial}{\partial t} \psi^* + H \psi^*) = 0$$

or, considering the operators p_i to be acting to the left and making use of (7.28),

$$\psi^* (-i\hbar \frac{\partial}{\partial t} + H) = 0.$$

Multiplying this equation on the right by ψ and making use of (7.27), we arrive at Equation (7.32).

Let us multiply Equation (7.32) on the right by ψ , and Equation (7.30) on the left by ψ^* . Then adding these we obtain

$$\frac{\partial}{\partial t} (\psi^* \psi) + \nabla \cdot (\mathbf{p} \psi^* \psi) = 0.$$

(The operator p acts on the function on its right.) Since

$$p_j = -i\hbar \frac{\partial}{\partial x_j} \quad (x_j = ix_j),$$

the above equation can be written

$$\frac{\partial}{\partial t} (\psi^* \psi) + \nabla \cdot (\mathbf{p} \psi^* \psi) = 0. \quad (7.33)$$

Equation (7.33) can be interpreted as a continuity equation. Writing it in the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0,$$

we obtain expressions for the density ρ and the current $\mathbf{j}$, namely

$$\left. \begin{aligned} \rho &= \psi^* \psi \\ \mathbf{j} &= \frac{i\hbar}{2m} (\psi^* \nabla \psi - \nabla \psi^* \psi) \end{aligned} \right\} \quad (7.34)$$

6. The Transformation Characteristics of Bilinear Combinations of the Wave Functions.

The γ_{ik} matrices can be used to give a unified form to the transformation equations (7.24) and (7.25), namely

$$\psi = (1 - \Delta_{ik} \gamma_{ik}) \psi', \quad (7.35)$$

where Δ_{ik} is the four-dimensional antisymmetric infinitesimal rotation tensor

$$\Delta_{xy} = -\frac{1}{4} \delta_{xy}, \quad \Delta_{4x} = -\frac{i}{4} (\delta_{0x})_x.$$

The equation which is the complex conjugate of (7.35) is

$$\psi^* = \psi'^* (1 - \Delta_{ik}^* \gamma_{ik}^*).$$

Multiplying this on the right by ψ , we obtain

$$\bar{\psi} = \bar{\psi}' (1 - \Delta_{ik} \gamma_{ik}). \quad (7.36)$$

These transformation equations can be used to prove that the product $\bar{\psi} \psi$ is a four-dimensional scalar. Indeed, according to (7.35) and (7.36)

$$\bar{\psi} \psi = \bar{\psi}' (1 - \Delta_{ik} \gamma_{ik}) (1 - \Delta_{ik}^* \gamma_{ik}^*) \psi',$$

or up to a term linear in Δ_{ik}

$$\bar{\psi} \psi = \bar{\psi}' \psi' - \Delta_{ik} \bar{\psi}' (\gamma_{ik} + \gamma_{ik}^*) \psi'.$$

But since Δ_{ik} is an antisymmetric tensor and $\gamma_{ik} + \gamma_{ik}^* = 2\delta_{ik}$, we have

$$\bar{\psi} \psi = \bar{\psi}' \psi'. \quad (7.37)$$

In a similar way we can prove that $\bar{\psi} \gamma_{ik} \psi$ is a four-dimensional vector. In fact, according to (7.35) and (7.36)

$$\bar{\psi} \gamma_i \psi = \bar{\psi}' \gamma_i \psi' - \Delta_{jk} \bar{\psi}' (\gamma_k \gamma_j \gamma_i + \gamma_i \gamma_j \gamma_k) \psi'$$

or, by making use of the properties of the γ matrices given in Equation (7.27) and the antisymmetry of Δ_{jk} , we obtain

$$\bar{\psi} \gamma_i \psi = \bar{\psi}' \gamma_i \psi' - 4 \Delta_{ik} \bar{\psi}' \gamma_k \psi'. \quad (7.38)$$

But (7.38) is exactly the expression of the transformation properties of a four-dimensional vector under infinitesimal rotation.

Similarly, it is easy to show that

$$\bar{\psi} \gamma_i \gamma_k \psi \quad (i \neq k)$$

is a tensor of second rank, that

$$\bar{\psi} \gamma_i \gamma_j \gamma_k \psi \quad (i \neq j \neq k)$$

is a tensor of third rank, and that

$$\bar{\psi} \gamma_i \gamma_k \gamma_j \gamma_m \psi \quad (i \neq k \neq j \neq m)$$

is a tensor of fourth rank.

Due to the commutation rules of the γ_i , it is impossible to construct higher rank tensors out of products of ψ and $\bar{\psi}$. We note that the above four-dimensional tensor of fourth rank is equivalent to the pseudoscalar

$$\bar{\psi} \gamma_5 \psi,$$

and that the third rank tensor is equivalent to the pseudovector

$$\bar{\psi} \gamma_5 \gamma_i \psi,$$

where

$$\gamma_5 = \gamma_1 \gamma_2 \gamma_3 \gamma_4.$$

(7.39)

8. Electron and Positron States.

1. Solutions with Positive and Negative Frequencies.

The general solution of the Dirac equation (7.11) can be written in the form of a Fourier integral

$$\psi = \int \psi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{k}. \quad (8.1)$$

The Fourier transform $\psi(\mathbf{k})$ is the wave function of the electron in momentum space. Inserting (8.1) into (7.11), we obtain

$$i \frac{\partial \psi(\mathbf{k})}{\partial t} = (\alpha \mathbf{k} + \beta m) \psi(\mathbf{k}); \quad (8.2)$$

or, writing $\psi(\mathbf{k})$ in terms of its component spinors,

$$\psi(\mathbf{k}) = \begin{pmatrix} \varphi(\mathbf{k}) \\ \chi(\mathbf{k}) \end{pmatrix},$$

This can be written in the form

$$\left. \begin{aligned} i \frac{\partial \varphi(\mathbf{k})}{\partial t} &= m \varphi(\mathbf{k}) + \sigma \mathbf{k} \chi(\mathbf{k}), \\ i \frac{\partial \chi(\mathbf{k})}{\partial t} &= -m \chi(\mathbf{k}) + \sigma \mathbf{k} \varphi(\mathbf{k}). \end{aligned} \right\} \quad (8.3)$$

We shall attempt to find solutions to this system in the form

$$\left. \begin{aligned} \varphi(\mathbf{k}) &= \varphi_0(\mathbf{k}) e^{-i\omega t}, \\ \chi(\mathbf{k}) &= \chi_0(\mathbf{k}) e^{-i\omega t}, \end{aligned} \right\} \quad (8.4)$$

so that after this is substituted into (8.3) we obtain

$$(\alpha \mathbf{k} + \beta m - \omega) \psi(\mathbf{k}) = 0, \quad (8.5)$$

function of k , and the other is expressed in terms of the first by means of Equation (8.6).¹⁾ For instance, if $\varphi^{(+)}$ and $\chi^{(-)}$ are given, we obtain $\varphi^{(-)}$ and $\chi^{(+)}$ as follows:

$$\left. \begin{aligned} \chi^{(+)} &= \frac{ek}{\epsilon + m} \varphi^{(+)}, \\ \varphi^{(-)} &= -\frac{k}{\epsilon + m} \chi^{(-)}. \end{aligned} \right\} \quad (8.12)$$

Solutions with positive and negative frequencies belong to different eigenvalues of the self-conjugate operator H (7.12). Thus, they are mutually orthogonal,²⁾

$$\psi^{(+)*}(\mathbf{k}) \psi^{(-)}(\mathbf{k}) = \varphi^{(+)*}(\mathbf{k}) \varphi^{(-)}(\mathbf{k}) + \chi^{(+)*}(\mathbf{k}) \chi^{(-)}(\mathbf{k}) = 0, \quad (8.13)$$

or

$$\int \psi^{(+)*}(\mathbf{r}) \psi^{(-)}(\mathbf{r}) d\mathbf{r} = 0. \quad (8.14)$$

Equation (8.13) is easy to prove with the aid of (8.12).

We note that the resolution (8.10) is relativistically invariant. Under Lorentz transformations the sign of the frequency will not change. This can be seen from the fact that the lowest positive frequency is m , and the highest negative frequency is $-m$, so that these are separated by an interval of $2m$, whereas the Lorentz transformations contain a continuous parameter.

2. The Wave Function of the Position.

The occurrence of two kinds of solutions has a fundamental meaning for the theory of the electron. Let us consider the particular solution of the Dirac equation in the form of a monochromatic wave

$$\psi = \psi_0(\mathbf{r}) e^{-i\omega t}. \quad (8.15)$$

Inserting (8.15) into (7.11), we find that $\psi_0(\mathbf{r})$ is an eigenfunction of the Hamiltonian, namely

¹⁾ The fact that not one but two components are arbitrary, is due to the form of the determinant of (8.6). Instead of considering a fourth order determinant, we are dealing with the second order determinant (8.7), which contains the matrices α . The fact that this determinant is independent of these matrices, shows that when it vanishes, so do its third order minors.

²⁾ The symbol $\varphi^* \varphi$ represents the scalar product of spinors given by

$$\varphi^* \varphi = \sum_{k=1}^4 \varphi_k^* \varphi_k.$$

$$H\psi_0 = \omega\psi_0 = \eta e\psi_0. \quad (8.16)$$

In quantum mechanics the eigenvalues of the Hamiltonian are the particle energies. In the relativistic quantum mechanics of the electron it is impossible to maintain this interpretation literally. Indeed, the occurrence of negative frequencies would mean both that there exist electron states with negative energies, and that there exists no lowest energy state. This would mean that in interacting with other particles the electron could transfer unlimited energy, going over into lower and lower energy states. The physical absurdity of this conclusion makes it necessary to change the fundamental quantum mechanical rules for calculating a physical quantity from the known solutions of the corresponding wave equation.

The new rules should, first of all, assign positive energies to negative frequency states. We shall assume that the electron energy for both signs of the frequency is given by

$$e = |\omega|.$$

In this way, however, there correspond to a given energy value (as well as to the values of other physical quantities which characterize completely the dynamic state of the electron, for instance, the momentum and spin orientation or angular momentum and parity) two different states, namely the states with different signs of their frequencies (quantum number η).

Therefore, the new rules should, in addition, give some physical meaning to the quantum number η . Such meaning can be obtained from the assumption that the Dirac equation refers to electrons with both positive and negative charge, or to the electron and the positron. It then becomes possible to interpret the quantum number η as characterizing the charge state of the electron. When $\eta = 1$ the electron has a charge e , and when $\eta = -1$ it has a charge $-e$. We shall see later that this hypothesis makes it possible to construct a theory of electrons and positrons and their interaction with the electromagnetic field.

In order to formulate the new quantum mechanical rules in correspondence with the above demands, let us establish the relation between solutions to the Dirac equation which have different signs for their frequencies. The wave function with negative frequency, according to (8.5) satisfies the equation

$$(\alpha k + \beta m + e)\psi^{(-)}(k) = 0.$$

Let us consider the complex conjugate equation

$$(\alpha^* k + \beta^* m + e)\psi^{(-)*}(k) = 0. \quad (8.17)$$

Let us multiply Equation (8.17) on the left by the matrix

$$C = i\beta\alpha_y = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (8.18)$$

This matrix, as can be seen from its definition, has the following properties:

$$\left. \begin{aligned} C &= \tilde{C} = C^*, \quad C^2 = 1, \\ C\beta &= -\beta C, \\ C\alpha^* &= \alpha C. \end{aligned} \right\} \quad (8.19)$$

(We note that $\beta^* = \beta$; $\alpha^* x = \alpha_x$; $\alpha^* y = -\alpha_y$.)

Equations (8.19) can be used to write (8.17) in the form

$$(-\alpha k + \beta m - e)C\psi^{(-)*}(k) = 0. \quad (8.20)$$

Let us now define that function

$$\psi^{(1)}(k) = C\psi^{(-)*}(-k). \quad (8.21)$$

Then according to (8.19)

$$\psi^{(1)}(-k) = C\psi^{(1)*}(k). \quad (8.22)$$

From (8.20) and (8.22) it follows that $\psi^{(1)}(k)$ satisfies the equation

$$(\alpha k + \beta m - e)\psi^{(1)}(k) = 0. \quad (8.23)$$

Going over from the momentum-space wave function $\psi^{(1)}(k)$ to the configuration-space wave function $\psi^{(1)}(r)$ by means of the Fourier transformation

$$\psi^{(1)}(r) = \int \psi^{(1)}(k) e^{ikr} dk,$$

we find that the function

$$\psi^{(1)}(r) = C\psi^{(-)*}(r) \quad (8.24)$$

satisfies the Dirac equation for a positive-frequency wave function.

Thus, if we consider the electron wave function with $\eta = -1$, not the function $\psi(-)$, but the function $\psi(p)$, then the Hamiltonian H may still be considered the energy operator. Its eigenvalues are now positive. We may call $\psi(p)$ the positron wave function.³⁾ The assignment of a charge whose sign is opposite to the electron charge to the state with $\eta = -1$ is merely a formal procedure so long as we restrict our considerations to the free electron and do not consider its interaction with the electromagnetic field. Later, however, we shall see (see Sections 11 and 19), that when the transformation (8.24) is applied to the equations of motion of the electron in the electromagnetic field it is found that the state whose wave function is $\psi(p)$ actually corresponds to the charge $+e$.

3. Positron Parity.

Let us now go on to a consideration of what happens to the operators for the other (in addition to the energy) physical quantities, for the electron in a charge state with $\eta = -1$. In doing this we can make use of the fact that the fundamental quantities, the momentum, the angular momentum, and the parity, are related to definite transformations of the wave function, in particular, to translations, rotations, and reflections. Let us first establish the form of the transformation operator for the function (8.24), and then let us go in the usual way to the operator corresponding to the physical quantity. Let the transformation for the function ψ be of the form.

$$\psi = K\psi'.$$

Then the function $\psi(p)$ will transform according to

$$\psi(p) = C\psi(-)^* = CK^*\psi'(-)^* = CK^*C\psi'p'.$$

Thus, the transformation operator for the positron function is

$$K(p) = CK^*C. \quad (8.25)$$

It follows from (8.25) that the momentum and angular momentum operators are of the same form for the positron and the electron. In particular, the infinitesimal translation operator is

$$\psi = (1 + i\mathbf{p})\psi',$$

or

$$K = 1 + i\mathbf{p}p,$$

³⁾ We shall sometimes denote the positive-frequency wave function $\psi(+)$ by the symbol $\psi(e)$.

⁴⁾ The transformation (8.21) is not the only one for which the equation $H\psi = -\omega\psi$ is transformed into $H\psi' = \omega\psi'$. The transformation $\psi \rightarrow p\psi$ will also perform this function. But if we demand, in addition, that the sign of the charge changes in the equation of motion of the electron in the electromagnetic field [transition from Equation (11.1) to (11.4)], then (8.21) is the only transformation possible:

where $\mathbf{p}$ is the momentum operator and $\mathbf{p}$ is the translation vector. Since

$$\mathbf{p} = -\mathbf{p}^*$$

and

$$CC^* = C^2 = 1,$$

it follows from (8.25) that

$$K(p) = K$$

and, therefore,

$$\mathbf{p}(p) = \mathbf{p}.$$

Similarly, the infinitesimal rotation operator is given by

$$\psi = (1 + \mathbf{J}\delta)\psi',$$

where δ is the infinitesimal rotation vector, and $\mathbf{J}$ is the rotation operator related to the angular momentum operator $\mathbf{M}$ by

$$\mathbf{J} = i\mathbf{M}.$$

The angular momentum operator of the electron consists of two terms,

$$\mathbf{M} = \mathbf{L} + \frac{1}{2}\boldsymbol{\Sigma}, \quad (8.26)$$

where $\mathbf{L}$ is the orbital angular momentum

$$\mathbf{L} = [\mathbf{r}\mathbf{p}],$$

and $\frac{1}{2}\boldsymbol{\Sigma}$ is the spin operator. Since

$$L^* = -L, \quad C\Sigma^*C = -\Sigma,$$

we have in this case

$$J(p) = J$$

and, therefore,

$$M(p) = M.$$

We obtain a different result for the inversion operator.¹⁾ According to (7.22) the inversion operator is

$$I = I_p \beta, \quad (8.27)$$

where I_p is the operator which changes the sign of the coordinates according to

$$I_p \psi(r) = \psi(-r).$$

Since according to (8.19)

$$C\beta C = -\beta,$$

it follows from (8.25) that the inversion operator $I(p)$ differs from the electron inversion operator I in the following way:

$$I(p) = -I_p I = -I. \quad (8.28)$$

In order to investigate the meaning of (8.28) we need only compare the electron and positron states whose wave functions coincide. Then the energy, momentum, and angular momentum of both particles are the same, whereas the parities are opposite.²⁾

¹⁾ V. Berestetsky, J. Exptl.-Theoret. Phys. 21, 43 (1951).

²⁾ We note that if the inversion properties of spinors are defined according to (7.8'), that is if $I = I_p \beta$, then $I(p) = I$. This has no effect on physical properties related to the parity conservation law. The parity operator of the electron-positron system $\Pi(p)$ is the same for both definitions.

4. Charge-Conjugate Function.

It should be emphasized that in defining the positron wave function according to Equation (8.24) we are not actually remaining within the realm of permissible quantum mechanical transformations. Part of the solution of the Dirac equation is subjected to a nonlinear transformation (the transition to the complex conjugate is a nonlinear operation). As a result the general solution to the Dirac equation, which can now be written

$$\psi = \psi^{(e)} + C\psi^{(p)*} \quad (8.29)$$

cannot be thought of as a wave function. Furthermore, a superposition of states with opposite signs of their charges, such as

$$\psi^{(e)} + \psi^{(p)}$$

does not form a general solution of the Dirac equation.³⁾ This situation, of course, is of no significance so long as we are considering free electron states with a definite charge. In the following we shall see also that the necessity for using the transformation of (8.24) does not interfere with the development of a general theory of electron systems and their interaction with the electromagnetic field.

We note also that (8.24) introduces only an apparent asymmetry in our method for considering electron and positron states. Let us consider the following transformation of a general solution to the Dirac equation:

$$\psi' = C\psi^*$$

or in other words

$$\psi' = \bar{\psi} C', \quad (8.30)$$

where

$$C' = \beta C. \quad (8.31)$$

It is easy to see that ψ' also satisfies the Dirac equation (7.30). The solution (8.29) transforms according to

$$\psi' = \psi^{(p)} + C\psi^{(e)*}.$$

Thus, ψ' differs from ψ by an interchange of the electron and positron wave functions. The function ψ' is called the solution which is charge-conjugate to ψ . In Section 19 we shall show the invariance of the theory with respect to the transformation (8.30) under more general conditions.

³⁾ This is true because a solution with positive frequencies is not a complete set of functions.

§ 9. Momentum and Polarization Eigenstates.

1. Plane Waves.

In this section and the one following we shall consider electron states with a given sign of the charge and with a given energy. Then the wave function of the electron is a solution of the equation

$$H\psi = e\psi, \quad (9.1)$$

where H is the Hamiltonian given by (7.12). Here ψ is an electron wave function [$\psi(e)$] or a positron wave function [$\psi(p)$]. We shall drop the time-dependent factor, since it is always of the form

$$e^{-iEt}.$$

States with definite values of η and ϵ are degenerate. We can distinguish these states by demanding that their wave functions be eigenfunctions also of other operators which commute with each other and with the Hamiltonian.

Let us start with a consideration of momentum eigenstates. The corresponding wave functions shall be denoted by ψ_p . They satisfy the equation

$$p\psi_p = p\psi_p,$$

whence

$$\psi_p = A e^{i p r}, \quad (9.2)$$

where A is a constant bispinor.

The wave function (9.2) should be normalized in some definite way. The normalization conditions can be chosen in various ways. If we consider the electron to be localized in some large volume V , then the normalization condition is

$$\int_V |\psi_p|^2 d\mathbf{r} = 1.$$

In this case

$$\psi_p = \frac{u}{\sqrt{V}} e^{i p r}, \quad (9.3)$$

where u is a unit bispinor satisfying the condition

$$u^\dagger u = 1. \quad (9.4)$$

Normalizing for an electron in an unbounded volume, according to the condition

$$\int |\psi_p|^2 d\mathbf{r} = 1$$

the momentum must be considered not exactly, but "almost" definite. With the aid of the Fourier transformation (8.1), this normalization condition can be written in the form

$$(2\pi)^3 \int |\psi_p(k)|^2 d\mathbf{k} = 1.$$

If $\psi_p(k)$ differs from zero only in a small (three-dimensional) region $d\mathbf{k}$ in the neighborhood of

$$k = p,$$

then

$$\psi_p(k) = \delta_{p,k} \frac{u}{(2\pi)^{3/2} \sqrt{d p}},$$

where u is the same unit bispinor as in (9.3) and (9.4). Inserting this expression in the Fourier integral (8.1), we obtain

$$\psi_p = \frac{\sqrt{d p}}{(2\pi)^{3/2}} u e^{i p r}. \quad (9.5)$$

The wave function (9.5) normalized "in an interval $d\mathbf{p}$ " differs by a factor $\sqrt{d p}$ from that normalized "in a unit interval of momentum", i.e., according to

$$\int \psi_p \psi_{p'} d\mathbf{r} = \delta(p - p').$$

2. Polarization States.

The bispinor u , in addition to the normalization condition (9.4), must satisfy Equation (8.3), or

$$(x p + \beta m - \epsilon) u = 0.$$

Let us write $\underline{u}$ in the form of a column vector consisting of two spinors $\underline{v}$ and $\underline{w}$:

$$\underline{u} = N \begin{pmatrix} \underline{v} \\ \underline{w} \end{pmatrix}, \quad (9.6)$$

where N is the normalizing constant. The spinors $\underline{v}$ and $\underline{w}$ are related by Equation (8.6). The spinor $\underline{v}$ may be considered arbitrary, and then $\underline{w}$ is given, according to (8.12), by

$$\underline{w} = \frac{\sigma p}{\epsilon + m} \underline{v}. \quad (9.7)$$

We shall choose the spinor $\underline{v}$ as a unit spinor, that is as one satisfying the condition

$$\underline{v}^* \underline{v} = 1. \quad (9.8)$$

From (9.6), (9.7), and (9.4) it follows that the constant N in (9.6) is given by

$$N = \frac{1}{\sqrt{1 + \frac{p^2}{(\epsilon + m)^2}}} = \sqrt{\frac{1}{2} \left(1 + \frac{m}{\epsilon} \right)}. \quad (9.9)$$

For a given momentum, there are two possible different electron states corresponding to two linearly independent two-component spinors $\underline{v}$. We shall speak of these two states as different polarization states.

Two linearly independent spinors $\underline{v}$ can be chosen, for instance, as eigenfunctions of the operator $\frac{1}{2} \sigma_z$. We shall consider the spinor index λ an independent variable which can take on the two values $\lambda = \pm \frac{1}{2}$. The eigenfunctions which belong to the eigenvalue μ of the operator $\frac{1}{2} \sigma_z$, shall be denoted $\underline{v}_\mu(\lambda)$. From the eigenvalue equation

$$\frac{\sigma_z}{2} \underline{v}_\mu(\lambda) = \mu \underline{v}_\mu(\lambda), \quad (9.10)$$

and expression (7.3) for the matrix σ_z , we find that μ can take on the values

$$\mu = \pm \frac{1}{2}$$

and that

$$\underline{v}_\mu(\lambda) = \delta_{\mu\lambda}. \quad (9.11)$$

Let us note that the second spinor $\underline{w}$ is not necessarily in this case an eigenfunction of the operator σ_z . Indeed, according to (9.7)

$$\underline{w}_\mu(\lambda) = \frac{1}{\epsilon + m} \sum_{\lambda'} (\sigma p)_{\lambda\lambda'} \underline{v}_\mu(\lambda') = \frac{1}{\epsilon + m} (\sigma p)_{\lambda\mu}.$$

If, however, we choose the $\underline{z}$ axis so that it is directed along the vector $\underline{p}$, then

$$\begin{aligned} \sigma p &= p \sigma_z, \\ (\sigma p)_{\lambda\mu} &= 0, \quad \lambda \neq \mu, \end{aligned}$$

and $\underline{w}$ is also an eigenfunction of σ_z , namely

$$\frac{1}{2} \sigma_z \underline{w}_\mu(\lambda) = \mu \underline{w}_\mu(\lambda) = \frac{p}{\epsilon + m} \underline{v}_\mu(\lambda).$$

These results follow from the fact that the electron spin operator

$$\frac{1}{2} \Sigma = \frac{1}{2} \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$$

does not commute with the Hamiltonian

$$H = \alpha \mathbf{p} + \beta m,$$

whereas the operator $\mathbf{p} \Sigma$ does commute with the Hamiltonian. Therefore, the eigenfunctions of H

$$\underline{u}_\mu = N \begin{pmatrix} \underline{v}_\mu \\ \underline{w}_\mu \end{pmatrix}$$

are not eigenfunctions of Σ_z , though they may be eigenfunctions of $\mathbf{p} \Sigma$, so that

$$\frac{1}{2} \frac{\Sigma \mathbf{p}}{p} \underline{u}_\mu = \mu \underline{u}_\mu. \quad (9.12)$$

States for which the bispinor u is defined by (9.10) and (9.11), or by (9.10) and (9.12) will be called spin or polarization eigenstates. In the latter case they are in addition eigenstates of the projection of the spin along the direction of motion. In both cases, as can be easily shown, the bispinors u_μ belonging to different values of μ are orthogonal:

$$u_\mu^* u_{\mu'} = \sum_\lambda u_\mu^*(\lambda) u_{\mu'}(\lambda) = \delta_{\mu\mu'}. \quad (9.13)$$

3. Sum Over Polarizations.

In the future we shall often come across the following problem. Given an expression $M_{\mu\nu}$, of the form

$$M_{\mu\nu} = \bar{u}_{\mu'} T u_\mu = \sum_{\lambda\lambda'} \bar{u}_{\mu'}(\lambda') T_{\lambda'\lambda} u_\mu(\lambda), \quad (9.14)$$

where T is some four-dimensional matrix (whose elements are $T_{\lambda'\lambda}$); the index ν takes on the values $\nu = 1, 2, 3, 4$, and is related to the spin index λ in the following way (see page 55).

$$u = \begin{pmatrix} u(1) \\ u(2) \\ u(3) \\ u(4) \end{pmatrix} = \begin{pmatrix} v(1/2) \\ v(-1/2) \\ w(1/2) \\ w(-1/2) \end{pmatrix};$$

the bispinors u and u' refer to different values of the momentum, namely p and p' , respectively; and $\bar{u}$ is related to u^* by Equation (7.31). We wish to obtain the sum

$$S = \sum_{\mu=\pm 1/2} \sum_{\mu'=\pm 1/2} |M_{\mu\mu'}|^2$$

of the squares of the absolute values of (9.14) over the different electron spin states.

We shall here present a convenient method for calculating such sums.

Since

$$\bar{u} = u^* \beta, \quad \bar{u}' T u = u'^* \beta T u,$$

the desired sum can be written in the form

$$S = \sum_{\mu\mu'} (\bar{u}_{\mu'} \beta T u_\mu) (u_\mu^* T^+ \beta u_{\mu'}) = \sum_{\mu\mu'} (\bar{u}_{\mu'} T u_\mu) (\bar{u}_\mu T^+ u_{\mu'}), \quad (9.15)$$

where

$$T^+ = \beta T^* \beta \quad (9.16)$$

and T^* is the Hermitian conjugate of T . Expression (9.15) contains products of the matrix elements of the operators βT and $T^+ \beta$ connecting the states μ and μ' , namely

$$S = \sum_{\mu\mu'} (\mu' | \beta T | \mu) (\mu | T^+ \beta | \mu').$$

The summation method consists of applying matrix multiplication rules to (9.15). However, direct use of these multiplication rules is impossible, since the bispinors u_μ entering the expression do not form a complete set. In fact, as can be seen from the Dirac equation in the form (7.30), u_μ satisfies the equation

$$(i\nabla p + m) u = 0,$$

where

$$\nabla p = \nabla p + \gamma_4 p_4, \quad p_4 = i\sqrt{p^2 + m^2}$$

that is, u_μ is an eigenfunction of the operator $i\nabla p$ belonging to the eigenvalue $-m$. We are considering two eigenfunctions u_μ of this operator ($\mu = \pm 1/2$). However, since the operator $i\nabla p$ is a four-dimensional matrix, it has not two but four eigenfunctions. The second pair belongs to the eigenvalue $+m$. Therefore, by considering the functions $u_\mu(\nu)$, which satisfy¹⁾

$$i\nabla p u_{\mu\Lambda} = -\Lambda u_{\mu\Lambda}, \quad (\Lambda = \pm m),$$

we obtain a system of functions $u_{\mu\Lambda}$ which is a complete system.

The functions $u_{\mu\Lambda}$ which enter into (9.15) have $\Lambda = m$. However, the sum can be extended also to $\Lambda = -m$ if we note that the operator

¹⁾ The eigenfunctions $u_{\mu\Lambda}$ with $\Lambda = -m$ are simply related to the bispinors $\bar{u}(-)$ corresponding to the negative-frequency solutions of the Dirac equation:

$$u_{\mu, -m} = \beta u^{\dagger(-)}$$

$$\frac{m - i\eta p}{2m}$$

is a projection operator; that is, it multiplies the function $\underline{u}_\mu \underline{m}$ by unity, and the function $\underline{u}_\mu - \underline{m}$ by zero:

$$\frac{m - i\eta p}{2\Lambda} \cdot u_{\mu\Lambda} = \delta_{\mu\Lambda} u_{\mu m}.$$

Thus,

$$S = \sum_{\mu\mu'} \sum_{\Lambda\Lambda'} \left(\bar{u}_{\mu'\Lambda'} \cdot \frac{m - i\eta p}{2\Lambda} u_{\mu\Lambda} \right) \left(\bar{u}_{\mu\Lambda} \cdot \frac{m - i\eta p'}{2\Lambda} u_{\mu'\Lambda'} \right). \quad (9.17)$$

Before making use of the completeness of the system $\underline{u}_{\mu\Lambda}$, let us exhibit the orthogonality and normalization properties of these functions. We shall prove the relation

$$\bar{u}_{\mu\Lambda} u_{\mu'\Lambda'} = \sum_{\nu} \bar{u}_{\mu\Lambda}(\nu) u_{\mu'\Lambda'}(\nu) = \frac{\Lambda}{\epsilon} \delta_{\mu\mu'} \delta_{\Lambda\Lambda'}. \quad (9.18)$$

For this purpose we note that the functions $\underline{u}_{\mu\Lambda}$ (the eigenfunctions of the non-Hermitian operator $\underline{i\eta p}$) are not orthogonal in the usual sense, i.e.,

$$u_{\mu\Lambda} u_{\mu'\Lambda'} \neq 0 \quad (\Lambda \neq \Lambda').$$

It is true, however, that

$$\bar{u}_{\mu\Lambda} u_{\mu'\Lambda'} = 0 \quad (\Lambda \neq \Lambda'),$$

which follows from the fact that according to (7.32) $\bar{u}_{\mu\Lambda}$ satisfies the equation

$$\bar{u}_{\mu\Lambda} (i\eta p - \Lambda) = 0.$$

To establish the orthogonality properties of the functions with different μ , we shall make use of the fact that the $\bar{u}_{\mu\Lambda}$ are eigenfunctions of the operator $\frac{1}{\Lambda} \Sigma p$, namely

$$\frac{1}{2} \Sigma p u_{\mu\Lambda} = \mu u_{\mu\Lambda}.$$

The complex conjugate of this equation is

$$u_{\mu\Lambda}^* \frac{1}{2} \Sigma p = \mu u_{\mu\Lambda}^*.$$

Multiplying this on the right by $\bar{u}_{\mu'}$, and noting that $\bar{u}_{\mu'}$ commutes with Σ , we obtain

$$\bar{u}_{\mu\Lambda} \frac{1}{2} \Sigma p = \mu \bar{u}_{\mu\Lambda}.$$

Thus, in addition to the usual orthogonality relations for functions with different values of μ , these functions also satisfy

$$\bar{u}_{\mu\Lambda} u_{\mu'\Lambda'} = 0 \quad (\mu \neq \mu').$$

Finally, in order to find the normalization of the functions $\underline{u}_{\mu\Lambda}$, we shall make use of expressions (9.6), (9.7), and (9.8) which are valid for all the $\underline{u}_{\mu\Lambda}$ if $\underline{m}$ is replaced by Λ . We then obtain

$$\bar{u}_{\mu\Lambda} u_{\mu\Lambda} = u_{\mu\Lambda}^* \bar{u}_{\mu\Lambda} = N^2 (\sigma_{\mu\Lambda}^* \sigma_{\mu\Lambda} - w_{\mu\Lambda}^* w_{\mu\Lambda}) = N^2 \left(1 - \frac{p^2}{(\epsilon + \Lambda)^2} \right),$$

whereas according to (9.4)

$$u_{\mu\Lambda}^* u_{\mu\Lambda} = N^2 (\sigma_{\mu\Lambda}^* \sigma_{\mu\Lambda} + w_{\mu\Lambda}^* w_{\mu\Lambda}) = N^2 \left(1 + \frac{p^2}{(\epsilon + \Lambda)^2} \right) = 1.$$

Therefore,

$$\bar{u}_{\mu\Lambda} u_{\mu\Lambda} = \frac{1 - \frac{p^2}{(\epsilon + \Lambda)^2}}{1 + \frac{p^2}{(\epsilon + \Lambda)^2}} = \frac{\Lambda}{\epsilon}.$$

Thus, we have proved Equation (9.18).

In view of the completeness of the set of $\bar{u}_{\mu\Lambda}$, it follows from (9.18) that

$$\sum_{\mu\Lambda} \frac{1}{\Lambda} \bar{u}_{\mu\Lambda}(\nu) u_{\mu\Lambda}(\nu) = \delta_{\nu\nu'}, \quad (9.19)$$

which shall be useful for our purposes.

Let us rewrite (9.17) in the form

$$S = \frac{1}{4} \sum_{\mu\Lambda} \sum_{\mu'\Lambda'} \bar{u}_{\mu\Lambda}(\nu) [T(m - i(p)) u_{\mu\Lambda}(\nu) \bar{u}_{\mu'\Lambda'}(\nu) \times \\ \times [\bar{T}(m - i(p')) u_{\mu'\Lambda'}(\nu) \frac{1}{\Lambda\Lambda'}].$$

Equation (9.19) can be used to sum over μ , μ' , and Λ , Λ' . We obtain as a result

$$S = \frac{1}{4\Lambda^2} \sum_{\nu, \nu'} [T(m - i(p)) u_{\nu}(\nu) \bar{T}(m - i(p')) u_{\nu'}(\nu')]$$

or,

$$S = \frac{1}{4\Lambda^2} \text{Sp} \{T(m - i(p)) \bar{T}(m - i(p'))\}, \quad (9.20)$$

where Sp denotes the trace (sum along the main diagonal) of the matrix.

4. Calculation of Traces.

Usually the matrix whose trace we want to calculate, as in (9.20), is a sum of terms each of which is a product of a number of γ_i ($i = 1, 2, 3, 4$). To calculate the trace we shall make use of the fact that

$$\text{Sp} \gamma_i \gamma_j \dots \gamma_n$$

is a four-dimensional tensor of rank n (where n is the number of factors). This follows from the fact that according to Section 7

$$\bar{u} \gamma_i \dots \gamma_n u$$

is a tensor of rank n , and therefore,

$$\sum_{\mu\Lambda} \bar{u}_{\mu\Lambda} \gamma_i \dots \gamma_n u_{\mu\Lambda} = \text{Sp} \gamma_i \dots \gamma_n.$$

is also a tensor. Since the γ_i have the same form in any coordinate system, the tensor $\text{Sp} \gamma_i \dots \gamma_n$ should also be independent of the coordinate system. The only tensor that has this property is δ_{jk} . Therefore, the desired tensor is composed of terms δ_{jk} .

Then it follows immediately that if n is an odd number

$$\text{Sp} \gamma_i \dots \gamma_n = 0 \quad (n = 2k + 1). \quad (9.21)$$

If n is an even number, then

$$\text{Sp} \gamma_i \dots \gamma_n = \sum_{\underline{j}} a_{\underline{j}} \delta_{i \underline{j} n},$$

where $\underline{j}, \underline{k}, \underline{l}, \underline{m}, \dots$ is some definite permutation of the indices $\underline{j}, \dots, \underline{n}$ and the $a_{\underline{j}}$ are to be found. The sum is taken over all possible pairs of combinations $\underline{j}, \underline{l}, \underline{m}, \dots$, and the number of terms in the sum, which is the number of such combinations, is clearly

$$\frac{n!}{2^{n/2} \left(\frac{n}{2}\right)!} = 1 \cdot 3 \cdot 5 \dots (n-1).$$

In order to determine the $a_{\underline{j}}$, all that is necessary is to give each pair of indices $\underline{j}, \underline{l}$, etc., the same values. Since

$$\gamma_i^2 = 1$$

and

$$\gamma_i \gamma_j = -\gamma_j \gamma_i \quad (i \neq j),$$

the trace will be multiplied by ± 1 , depending on whether it takes an even or an odd number of transpositions of the indices of the γ_i to bring the product

$$\gamma_i \gamma_k \gamma_l \dots$$

into the form

$$\gamma_{i_1} \dots \gamma_{i_n}$$

Further, since the trace of the unit matrix is equal to 4,

$$\text{Sp } 1 = 4,$$

it follows that

$$\frac{1}{4} \text{Sp } \gamma_{i_1} \gamma_{i_2} \dots \gamma_{i_n} = \sum_{\mathbf{P}} \delta_{i_1 i_2} \delta_{i_3 i_4} \dots \quad (n = 2k). \quad (9.22)$$

The rules for finding the coefficients $\frac{a_{\mathbf{P}}}{4}$ can be formulated in the following way.

Each matrix $\gamma_{\mathbf{j}}$ is made to correspond to a point $\mathbf{j}$ on a circle. The points on the circle are ordered in the same way as the matrices in the trace. Pairs of points are joined by straight lines. Then to each line joining points $\mathbf{j}$ and $\mathbf{k}$, corresponds a factor $\delta_{\mathbf{j}\mathbf{k}}$, and to each way of joining the points (or in other words to each way of breaking up the indices $\mathbf{j}, \mathbf{k}, \mathbf{l}, \mathbf{m}, \mathbf{n}, \mathbf{p}, \dots$ into pairs $\mathbf{j}-\mathbf{k}, \mathbf{l}-\mathbf{m}, \mathbf{n}-\mathbf{p}, \dots$) there corresponds a term

$(-1)^P \delta_{i_1 i_2} \delta_{i_3 i_4} \dots$ in the series for the trace, where P is the number of intersections of these lines. Thus,

$$\frac{a_{\mathbf{P}}}{4} = (-1)^P.$$

The values of $\text{Sp } \gamma_{\mathbf{j}_1} \dots \gamma_{\mathbf{j}_n}$ with $n = 2, 4, 6$ are

$$\left. \begin{aligned} \frac{1}{4} \text{Sp } \gamma_{\mathbf{i}\mathbf{k}} &= \delta_{\mathbf{i}\mathbf{k}}, \\ \frac{1}{4} \text{Sp } \gamma_{\mathbf{i}\mathbf{k}} \gamma_{\mathbf{l}\mathbf{m}} &= \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{l}\mathbf{m}} + \delta_{\mathbf{i}\mathbf{m}} \delta_{\mathbf{k}\mathbf{l}} - \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{m}}, \\ \frac{1}{4} \text{Sp } \gamma_{\mathbf{i}\mathbf{k}} \gamma_{\mathbf{l}\mathbf{m}} \gamma_{\mathbf{n}\mathbf{p}} &= \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{l}\mathbf{m}} \delta_{\mathbf{n}\mathbf{p}} + \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{l}\mathbf{p}} \delta_{\mathbf{m}\mathbf{n}} + \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{m}\mathbf{p}} \delta_{\mathbf{l}\mathbf{n}} + \delta_{\mathbf{i}\mathbf{m}} \delta_{\mathbf{k}\mathbf{l}} \delta_{\mathbf{n}\mathbf{p}} + \\ &+ \delta_{\mathbf{i}\mathbf{m}} \delta_{\mathbf{k}\mathbf{p}} \delta_{\mathbf{l}\mathbf{n}} + \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{m}} \delta_{\mathbf{n}\mathbf{p}} + \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{n}} \delta_{\mathbf{m}\mathbf{p}} - \\ &- \delta_{\mathbf{i}\mathbf{m}} \delta_{\mathbf{k}\mathbf{l}} \delta_{\mathbf{n}\mathbf{p}} - \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{l}\mathbf{p}} \delta_{\mathbf{m}\mathbf{n}} - \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{m}} \delta_{\mathbf{n}\mathbf{p}} - \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{n}} \delta_{\mathbf{m}\mathbf{p}} - \\ &- \delta_{\mathbf{i}\mathbf{k}} \delta_{\mathbf{l}\mathbf{m}} \delta_{\mathbf{n}\mathbf{p}} - \delta_{\mathbf{i}\mathbf{l}} \delta_{\mathbf{k}\mathbf{m}} \delta_{\mathbf{n}\mathbf{p}} - \delta_{\mathbf{i}\mathbf{m}} \delta_{\mathbf{k}\mathbf{l}} \delta_{\mathbf{n}\mathbf{p}} \end{aligned} \right\} \quad (9.23)$$

We note that the matrix $\gamma_{\mathbf{j}_1} \dots \gamma_{\mathbf{j}_n}$ can be simplified if among the indices $\mathbf{j}_\alpha$ there exist two which are equal, $\mathbf{j}_\alpha = \mathbf{j}_\beta = \mathbf{j}$ and the sum over them is taken. Equation (7.27) can be used to transform the product

$$\gamma_{\mathbf{j}_1} \dots \gamma_{\mathbf{j}_n}$$

so that the two matrices $\gamma_{\mathbf{j}}$ occur next to each other. Then, since

$$\sum_{\mathbf{j}} \gamma_{\mathbf{j}} \gamma_{\mathbf{j}} = 4,$$

the trace of this matrix is not a tensor of rank n , but one of rank $n-2$. For instance, if the two matrices $\gamma_{\mathbf{j}}$ are separated by one, two, or three factors, the following relations hold (summation over $\mathbf{j}$ from $\mathbf{j} = 1$ to $\mathbf{j} = 4$ is implied):

$$\left. \begin{aligned} \gamma_{\mathbf{j}} \gamma_{\mathbf{i}} \gamma_{\mathbf{j}} &= -2\gamma_{\mathbf{i}}, \\ \gamma_{\mathbf{j}} \gamma_{\mathbf{i}} \gamma_{\mathbf{k}} \gamma_{\mathbf{j}} &= 4\delta_{\mathbf{i}\mathbf{k}}, \\ \gamma_{\mathbf{j}} \gamma_{\mathbf{i}} \gamma_{\mathbf{k}} \gamma_{\mathbf{l}} \gamma_{\mathbf{j}} &= -2\gamma_{\mathbf{i}} \gamma_{\mathbf{k}} \gamma_{\mathbf{l}}. \end{aligned} \right\} \quad (9.24)$$

§ 10. Angular Momentum and Parity Eigenstates of the Electron.

1. Orbital and Spin Functions.

Let us go on to a consideration of eigenstates not only of the energy, but also of the angular momentum. The eigenfunctions of the square of the angular momentum $\mathbf{M}^2$ and of its projection M_z , belonging to the eigenvalues

$$\begin{aligned} M^2 &= J(J+1), \\ M_z &= M, \end{aligned}$$

shall be denoted by $\psi_{\mathbf{JM}}$. They satisfy the equations

$$\left. \begin{aligned} \mathbf{M}^2 \psi_{\mathbf{JM}} &= J(J+1) \psi_{\mathbf{JM}}, \\ M_z \psi_{\mathbf{JM}} &= M \psi_{\mathbf{JM}}, \end{aligned} \right\} \quad (10.1)$$

where $\mathbf{M}$ is the angular momentum operator given by (8.26).

As in Chapter I we shall not solve Equation (10.1) directly, but shall make use of the quantum mechanical rules for composition of angular momentum to find the desired wave functions.

We shall consider the two spinors making up the wave function $\psi_{\mathbf{JM}}$ separately. Let

$$\psi_{\mathbf{JM}} = \begin{pmatrix} \psi_{\mathbf{JM}} \\ \chi_{\mathbf{JM}} \end{pmatrix}.$$

Both φ_{JM} and χ_{JM} satisfy Equations (10.1) with the four-dimensional matrix Σ replaced by the two-dimensional one σ , i.e., with the angular momentum operator $\mathbf{M}$ written in the form

$$\mathbf{M} = \mathbf{L} + \frac{1}{2} \sigma.$$

The eigenfunctions of the orbital and spin angular momentum are known. We shall denote the eigenvalues of $\mathbf{L}^2$ and L_z as previously, by

$$\begin{aligned} \mathbf{L}^2 &= l(l+1), \\ L_z &= m \end{aligned}$$

and the orbital wave function by Φ_{lm} . This latter can be represented in the form

$$\Phi_{lm} = a(r) Y_{lm} \left(\frac{\mathbf{r}}{r} \right),$$

where $a(r)$ is a radial function which shall be determined later. The eigenfunctions $\chi_{\mu}(\lambda)$ of the spin angular momentum belonging to the eigenvalues of $\frac{1}{2} \sigma_z$ and $(\frac{1}{2} \sigma)^2$ equal to

$$\begin{aligned} \frac{1}{2} \sigma_z &= \mu, \\ \left(\frac{1}{2} \sigma \right)^2 &= \frac{1}{2} \left(\frac{1}{2} + 1 \right), \end{aligned}$$

were found in Section 9. According to (9.10) and (9.11)

$$\sigma_{\mu}(\lambda) = \delta_{\mu\lambda}.$$

We note that an arbitrary spinor φ can be expanded in a series of orthonormal spinors $\chi_{\mu}^{(j)}$ according to

$$\varphi(\lambda) = \sum_{\mu, j} \varphi_{\mu}^{(j)} \sigma_{\mu}(\lambda). \quad (10.2)$$

¹⁾ It may be said that the χ_{μ} form a "spinor basis" analogous to the χ_{μ} in Section 3.

We shall call the φ^{μ} the contravariant components of the spinor. Inserting the explicit expression for $\chi_{\mu}(\lambda)$, we find that

$$\varphi(\lambda) = \varphi^{\lambda},$$

which means that the contravariant components of the spinor defined by (10.2) are the same components as we have been using on the basis of the definition (7.1).

In addition to the contravariant spinor components we can make use also of the covariant components, which are defined by the condition that the scalar product of two spinors φ and η be of the form

$$\varphi\eta = \eta^{\mu}\varphi_{\mu}.$$

Since the scalar product of two spinors is given by the relation¹⁾

$$\varphi\eta = \varphi^{1/2}\eta^{1/2} - \varphi^{-1/2}\eta^{-1/2},$$

we have

$$\varphi_{\mu} = (-1)^{\mu-1/2} \varphi^{-\mu}. \quad (10.3)$$

2. Spherical Spinors.

Let us now return to the problem of determining φ_{JM} and χ_{JM} . We shall do this by the same method as we used for the photon in Section 4.

The function φ may be considered a scalar function in the generalized space of the coordinates $\mathbf{xyz}$ and spin variable λ of the electron. [Different spinor components $\varphi^{\lambda}(\mathbf{r})$ at a given point in position space correspond to values of the scalar function $\varphi(\mathbf{r}, \lambda)$ at different points of the generalized spin space.] Our problem reduces to finding the wave function of a system consisting of two noninteracting subsystems (the orbital and the spin degrees of freedom). The angular momentum of the spin subsystem is always $1/2$. Therefore, according to the rules for composition of angular momentum, a given total angular momentum $\mathbf{j}$ can be obtained only for two values of the orbital angular momentum

$$l = j \pm \frac{1}{2}.$$

The wave functions of two different states corresponding to the same values of $\mathbf{j}$ and $\mathbf{M}$ shall be denoted by

¹⁾ See, for instance, L. Landau and E. Lifshits, Quantum Mechanics (State Tech. Press, 1948) Section 54.

$$\varphi_{JM} \quad (l = j \pm \frac{1}{2})$$

These functions can be written in the form of a superposition of products of orbital and spin functions Φ_{lm} and χ_{μ} :

$$\varphi_{JM} = \sum_{m, \mu} C_{JM}^{lm, \mu} \Phi_{lm}(r) \chi_{\mu}(\lambda) = a(r) \sum_{m, \mu} C_{JM}^{lm, \mu} Y_{lm} Y_{1, M-\mu} \chi_{\mu}, \quad (10.4)$$

The coefficients entering into (10.4) differ from those in (4.3) by the fact that now $\lambda = 1/2$. Equation (10.4) can be written in the form

$$\varphi_{JM} = a(r) \Omega_{JM} \left(\frac{r}{r_0} \right), \quad (10.5)$$

where Ω_{JM} is a spinor whose contravariant components, according to (10.2), are given by

$$\Omega_{JM}^{\mu} = C_{JM}^{l, M-\mu, \mu} Y_{l, M-\mu} \chi_{\mu}. \quad (10.6)$$

The quantity Ω_{JM} shall be called a spinor spherical function or spherical spinor.¹⁾ The covariant components of a spherical spinor are, according to (10.3),

$$\Omega_{JM, \mu} = (-1)^{\mu} C_{JM}^{l, M+\mu, -\mu} Y_{l, M+\mu} \chi_{-\mu}.$$

We give here the values of the coefficients $C_{JM}^{lm, \mu}$ (we shall at times also use the notation $C_{JM}^{lm, \mu} = \frac{1}{2} \frac{1}{\sqrt{2l+1}}$):

$\mu \backslash l$	$l + \frac{1}{2}$	$l - \frac{1}{2}$
$\frac{1}{2}$	$\sqrt{\frac{l+M+\frac{1}{2}}{2l+1}}$	$-\sqrt{\frac{l-M+\frac{1}{2}}{2l+1}}$
$-\frac{1}{2}$	$\sqrt{\frac{l-M+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l+M+\frac{1}{2}}{2l+1}}$

(10.7)

¹⁾ For the definition of spinor spherical functions see also V. Fock, Introduction to Quantum Mechanics (Commission for Improving the Living Conditions of Students, 1952); V. Barstetsky, A. Dolgov, and K. Tor-Martirosyan, J. Exptl.-Theoret. Phys. 20, 327 (1955).

They are normalized so that

$$\sum_{\mu} (\chi_{\mu}^{\lambda})^2 = 1. \quad (10.8)$$

The spherical spinors are an orthonormal set of functions. In fact, for different values of J, M , or l , the functions Ω_{JM} are mutually orthogonal since they are eigenfunctions of Hermitian operators and belong to different eigenvalues. In view of (10.8) they are also normalized:

$$\int \Omega_{JM}^* \Omega_{J'M'} d\Omega = \delta_{JJ'} \delta_{MM'}. \quad (10.9)$$

3. Angular Momentum Eigenfunction.

In order to determine the radial dependence of the φ_{JM} , i.e., the function $a(r)$ in (10.5), we shall make use of the Fourier transformation for $\varphi(r)$:

$$\varphi(r) = \int \varphi(k) e^{ikr} dk.$$

To the angular momentum eigenfunctions there corresponds a momentum-space function $\varphi(k)$ which is an eigenfunction both of the square of the angular momentum and of its projection. The angular momentum operator in momentum space has the same structure as in configuration space, namely

$$M = L + \frac{1}{2} \sigma,$$

where σ has the previous meaning, and

$$L = -i[k \nabla_k].$$

We can therefore write immediately

$$\varphi_{JM}(k) = a(k) \Omega_{JM} \left(\frac{k}{k} \right). \quad (10.10)$$

The function $a(\mathbf{k})$, except for its normalization which we shall consider later, is determined by the fact that since the electron has a definite energy ϵ , it must also have a definite momentum $\mathbf{p}$ given by

$$\mathbf{p} = \sqrt{\epsilon^2 - m^2}.$$

We shall therefore assume that $a(\mathbf{k})$ differs from zero only in a small region $d\mathbf{p}$ in the neighborhood of $\mathbf{k} = \mathbf{p}$. Inserting (10.10) into the Fourier integral, we have

$$\varphi_{JM} = a(p) p^2 dp \int e^{i\mathbf{p}\cdot\mathbf{r}} \Omega_{JM}(\mathbf{n}) d\mathbf{n}.$$

Expanding $e^{i\mathbf{p}\cdot\mathbf{r}}$ in a series of spherical functions [see (4.28)], we obtain

$$\varphi_{JM} = a(p) p^2 dp \sum_{l'm'} g_{l'm'}(pr) Y_{l'm'}\left(\frac{\mathbf{r}}{r}\right) \int Y_{l'm'}^*(\mathbf{n}) \Omega_{JM}(\mathbf{n}) d\mathbf{n}.$$

If we insert the expression for the spherical spinors in terms of spherical functions [see (10.6)] into this expression, then it is easy to see that

$$Y_{l'm'}\left(\frac{\mathbf{r}}{r}\right) \int Y_{l'm'}^* \Omega_{JM} d\mathbf{n} = \delta_{l'l} \delta_{m'm} \Omega_{JM}^*\left(\frac{\mathbf{r}}{r}\right)$$

and, therefore,

$$\varphi_{JM} = a(p) p^2 dp g_l(pr) \Omega_{JM}. \quad (10.11)$$

Comparing (10.11) with (10.5), we see that

$$a(r) = a(p) p^2 dp g_l(pr).$$

To determine the second part χ_{JM} of the electron wave function, we make use of the fact that in momentum space $\chi(\mathbf{k})$ and $\varphi(\mathbf{k})$ are related simply by Equation (8.12), so that

$$\chi(\mathbf{k}) = \frac{mk}{\epsilon + m} \varphi(\mathbf{k}).$$

Since $\varphi_{JM}(\mathbf{k})$ is given by (10.10), we obtain

$$\chi_{JM}(\mathbf{k}) = \frac{a(k)k}{\epsilon + m} \left(\sigma \frac{\mathbf{k}}{k} \right) \Omega_{JM}\left(\frac{\mathbf{k}}{k}\right). \quad (10.12)$$

On the other hand according to (10.1) χ_{JM} , as well as φ_{JM} , is an eigenfunction of the operators $\mathbf{M}^2$ and M_z , and belongs to the same eigenvalues J and M . This means that its angular dependence is given by a spinor spherical function. From this we can conclude that the product

$$(\sigma \mathbf{n}) \Omega_{JM}(\mathbf{n})$$

which enters into (10.12) is a spherical spinor. Since according to (10.6) Ω_{JM} contains the spherical function Y_{JM} , it follows [see (4.13)] that

$$n Y_{JM}$$

contains the spherical function $Y_{J'M'}$, where $J' = J \pm 1$. Of these two values of J' , only one, namely

$$J' = 2J - J = \begin{cases} J+1 & \text{for } J = J - \frac{1}{2}, \\ J-1 & \text{for } J = J + \frac{1}{2}, \end{cases}$$

is compatible with the rules for composition of angular momenta with a given value of J . Thus, it follows that

$$\sigma n \Omega_{JM} = q \Omega_{J'M} \quad (J' = 2J - J, q = 1). \quad (10.13)$$

The validity of (10.13) is also easy to see with the aid of direct application of (10.6) and (4.13). Then the coefficient q is seen to be unity. In configuration space we obtain from (10.12) and (10.13) an expression for the χ_{JM} , which is analogous to the transition from (10.10) to (10.11):

$$\chi_{JM} = a(p) p^2 dp \frac{p}{\epsilon + m} g_{J'}(pr) \Omega_{J'M} \quad (J' = 2J - J). \quad (10.14)$$

4. Parity of a State.

Let us note that φ_{JM} and χ_{JM} belong to different eigenvalues of L^2 [the former to $l(l+1)$, and the latter to $l'(l'+1)$]. This means that the electron wave function

$$\psi_{JM} = \begin{pmatrix} \varphi_{JM} \\ \chi_{JM} \end{pmatrix}$$

is not an eigenfunction of the orbital angular momentum. It is similarly easy to see that L^2 does not commute with the Hamiltonian H . We thus arrive at the conclusion that the separation of the angular momentum into its orbital and spin parts has only limited physical meaning for the electron, as for the photon. This separation, however, becomes valuable in the nonrelativistic limit. As is seen, for instance, from (10.12), the ratio of χ to φ goes to zero as $\hbar \rightarrow 0$. This means that for low energies we may make use of a wave function consisting only of two components (since the other two are small). Then l takes on the meaning of the orbital angular momentum. In the general case, however, the index l on the wave function serves only to denote the two different electron states with the same values of J and M .

The quantum number l can be uniquely related to the parity of the state. Since the inversion operator, according to (8.27) and (8.28), is different for the electron and positron, we shall first consider electron states ($\eta = +1$).

Let us apply the operator of (8.27) to the wave function ψ_{JM} . Since

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where 1 denotes the two-dimensional unit matrix, we have

$$l\psi_{JM} = l\beta \begin{pmatrix} \varphi_{JM} \\ \chi_{JM} \end{pmatrix} = \begin{pmatrix} l\varphi_{JM} \\ -l\chi_{JM} \end{pmatrix}.$$

Further, since φ_{JM} contains the spherical function Y_{lm} , and χ_{JM} the function $Y_{l'm}$, and since

$$l_r Y_{lm} = (-1)^l Y_{lm}, \\ l_r Y_{l'm} = (-1)^{l'} Y_{l'm} = (-1)^{l'+1} Y_{l'm},$$

we arrive at⁹

⁹ We note that because the factor β appears in the inversion operator, the wave functions ψ_{JM} would not be eigenfunctions of the operator l if φ_{JM} and χ_{JM} belonged to the same value l .

$$l\psi_{JM} = (-1)^l \psi_{JM}. \quad (10.15)$$

Equation (10.15) determines the parity of the state in terms of the quantum number l . Since the two values l for a given J differ by unity, of the two electron states belonging to the same values of J and M one is even and the other is odd.

For positron states, it follows from (8.28) that

$$l\psi_{JM} = (-1)^{l+1} \psi_{JM}. \quad (10.16)$$

In the expression we obtained for the wave function [see (10.11), (10.14)] the factor $a(p)$ has so far remained unknown. It can be determined from the normalization conditions, which can be written

$$\int |\psi_{JM}|^2 d\mathbf{r} = (2\pi)^3 \int |\psi_{JM}(\mathbf{k})|^2 d\mathbf{k} = 1,$$

or, using (10.10) and (10.12),

$$(2\pi)^3 p^3 dp |a(p)|^2 \left[1 + \frac{p^2}{(s+m)^2} \right] = 1.$$

From this we obtain

$$a(p) = \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{1}{2} \left(1 + \frac{m^2}{s^2} \right)} \frac{1}{\sqrt{p^2 dp}}. \quad (10.17)$$

Inserting (10.17) into (10.11) and (10.14), we finally obtain an expression for the wave function of the electron which is an eigenfunction of the angular momentum, parity, and energy (in order to emphasize the latter, we shall provide the function with a fourth quantum number ϵ), namely

$$\left. \begin{aligned} \varphi_{JM} &= \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{(s+m)p d\epsilon}{2}} g_I(p\mathbf{r}) \Omega_{JM}, \\ \chi_{JM} &= \frac{1}{(2\pi)^{3/2}} \sqrt{\frac{(s-m)p d\epsilon}{2}} g_{II}(p\mathbf{r}) \Omega_{JM} \quad (l' = 2J - l). \end{aligned} \right\} \quad (10.18)$$

We here note the expressions for the wave functions for other normalization requirements. If the normalization conditions are such that

$$\int \psi_{JM}^* \psi_{JM} d\tau = \delta(\epsilon - \epsilon'),$$

then the wave function ψ_{JM} differs from (10.18) only by the absence of the factor $\sqrt{d\epsilon}$. If the wave function is normalized for a sphere of radius R , $d\epsilon$ should be replaced by $\frac{\pi R^3}{\epsilon}$, or $d\epsilon$ by $\frac{\pi R^3}{\epsilon^2}$, in (10.18).

5. Expansion in Spherical Waves.

We have now constructed two different complete sets of functions ψ_{JM} and ψ_{JM} . Any solution of the Dirac equation (with a given sign of its frequency) can be expanded in either one of these sets. Let $\psi(\mathbf{k})$ be the wave function of an electron in momentum space. Then

$$\psi(\mathbf{k}) = \sum_{\mathbf{n}} C_{\mathbf{n}} \psi_{\mathbf{n}}(\mathbf{k}), \quad (10.19)$$

where $\mathbf{n}$ represents the set of quantum numbers $\mathbf{p}$ or ϵ and $\mathbf{M}$.

If the $\psi_{\mathbf{n}}(\mathbf{k})$ are normalized so that

$$(2\pi)^3 \int \psi_{\mathbf{n}}^*(\mathbf{k}) \psi_{\mathbf{n}'}(\mathbf{k}) d\mathbf{k} = \delta_{\mathbf{n}\mathbf{n}'},$$

we obtain

$$C_{\mathbf{n}} = (2\pi)^3 \int \psi(\mathbf{k}) \psi_{\mathbf{n}}^*(\mathbf{k}) d\mathbf{k}. \quad (10.20)$$

In particular, we can expand a momentum and polarization eigenstate in angular momentum and parity eigenfunctions, obtaining

$$\psi_{JM} = \sum_{\mathbf{n}} C_{\mathbf{n}} \psi_{JM},$$

or the inverse expansion

$$\psi_{JM} = \sum_{\mathbf{n}} C_{JM}^{\mathbf{n}} \psi_{\mathbf{n}}.$$

According to (10.20), (10.18), and (9.5)

$$C_{JM}^{\mathbf{n}} = (C_{JM}^{\mathbf{n}})^* = \frac{1}{2} \left(1 + \frac{m}{\epsilon} \right) \left[v_{\mathbf{n}} \Omega_{JM} \left(\frac{p}{\epsilon} \right) + \frac{p}{\epsilon + m} w_{\mathbf{n}} \Omega_{JM} \left(\frac{p}{\epsilon} \right) \right] \sqrt{d\epsilon},$$

or since

$$w_{\mathbf{n}} = \frac{\sigma p}{\epsilon + m} v_{\mathbf{n}}$$

and

$$\Omega_{JM} = \left(\sigma \frac{p}{\epsilon} \right) \Omega_{JM},$$

we have

$$C_{JM}^{\mathbf{n}} = (C_{JM}^{\mathbf{n}})^* = \left(v_{\mathbf{n}} \Omega_{JM} \left(\frac{p}{\epsilon} \right) \right) \sqrt{d\epsilon}. \quad (10.21)$$

The quantities $|C_{JM}^{\mathbf{n}}|^2$ determine the angular distribution of electrons in angular momentum eigenstates. If we sum this quantity over spins, we obtain

$$\sum_{\mathbf{n}} |C_{JM}^{\mathbf{n}}|^2 = |\Omega_{JM}|^2 d\epsilon. \quad (10.22)$$

For a given J , equation (10.22) does not depend on M , since

$$\Omega_{JM}^* \Omega_{JM} = (\Omega_{JM}^* \sigma n) (\sigma n \Omega_{JM}) = |\Omega_{JM}|^2. \quad (10.23)$$

An expansion of the form (10.19) is valid also for configuration-space wave functions, so that

$$\psi(\mathbf{r}) = \sum_{\mathbf{n}} C_{\mathbf{n}} \psi_{\mathbf{n}}(\mathbf{r}),$$

where the coefficients C_{μ}^{ν} are given by (10.20) as previously. In particular, using (10.21) and the explicit expressions for ψ_{μ}^{ν} and ψ_{μ}^{ν} , we obtain the following expansion for an electron plane wave in terms of spherical waves:

$$ue^{i\mathbf{p}\cdot\mathbf{r}} = \sum_{\mu} \left(\psi_{\mu}^{\nu} \left(\frac{\mathbf{p}}{p} \right) \frac{1}{\sqrt{2}} \left(\frac{V}{1 + \frac{m}{E}} g_{\mu}^{\nu}(pr) \Omega_{\mu}^{\nu} \left(\frac{\mathbf{r}}{r} \right) \right) \right), (l' = 2j - l), \quad (10.24)$$

§ 11. The Electron in an External Field.

1. Dirac Equation in an External Field.

The theory of the interaction between the electron and the electromagnetic field will be developed in later chapters. This interaction leads in general to the creation and annihilation of photons and electron-positron pairs. Within the framework of the single-electron theory, however, which we are now considering, we can treat a more limited type of problem. In this type of problem the number of particles does not change and the interaction can be introduced on the basis of the external field concept.

The equation of motion of an electron in a given external field is easily obtained in the same way as it is done in nonrelativistic quantum mechanics. Let Δ be the four-dimensional potential of the external electromagnetic field (Δ is the vector potential, and $\Delta_0 = -i\Delta_4$ is the scalar potential). We obtain the desired equation if we replace the four-dimensional momentum operator p by $p - e\Delta$ in the Dirac equation:

$$\mathcal{D} \rightarrow \mathcal{D} - e\Delta.$$

Obviously, the equation remains relativistically invariant, since the transformation properties of p and Δ are the same.

Thus, the Dirac equation in an external field becomes

$$(\gamma p - ie\gamma A + m)\psi = 0. \quad (11.1)$$

Let us restrict our considerations to the case in which the external field is time independent. Then there exist stationary solutions of the form

$$\psi(\mathbf{r}, t) = \psi_0(\mathbf{r}) e^{-i\omega t},$$

where $\psi_0(\mathbf{r})$ is an eigenfunction of the Hamiltonian, i.e.,

$$\left. \begin{aligned} H\psi_0 &= \omega\psi_0, \\ H &= \alpha p + \beta m + eA_0 - e\mathbf{A} \cdot \mathbf{A}. \end{aligned} \right\} \quad (11.2)$$

A general solution of (11.1) may be represented as a superposition of wave functions with various frequencies ω .

There is a significant difference between the values of the frequencies in this case and in the free-electron case. In the absence of an external field, as we have seen in Section 8, we obtain a continuous frequency spectrum with a discontinuity between $-m$ and m , namely

$$m \leq \omega \leq \infty,$$

or

$$-\infty \leq \omega \leq -m.$$

In the presence of an external field, in addition to the continuous spectrum there may exist also a discrete spectrum (bound states) in the region¹⁾

$$-m < \omega < m.$$

We have interpreted the solutions with

$$\omega > m$$

as electron-state wave functions, and solutions with

$$\omega < -m$$

as positron-state functions transformed according to (8.24). There arises the question of interpreting the discrete spectrum. We shall consider it in more detail somewhat later, restricting ourselves for the present to that important case in which the discrete spectrum is located close to the boundary of one of the continuous spectra, i.e.,

$$|m - \omega_{\text{min}}| < m.$$

We can divide the frequencies into positive and negative ones, as previously, and interpret them according to the rules formulated in Section 8.

If a solution $\psi = \psi^{(-)}$ of Equation (11.1) belongs to negative frequencies, it is necessary to go over from this solution to a positron wave function $\psi^{(p)}$. Let us transform Equation (11.1) for this case.

¹⁾ See paragraph 4 of this section.

The complex conjugate of Equation (11.1) is

$$[-i\gamma^* (p^* - eA^*) - i\gamma_4^* (p_4^* - eA_4^*) + m]\psi^* = 0$$

which we shall multiply on the left by C . Since [see (8.19) and (7.26)]

$$\left. \begin{aligned} C\gamma^* &= \gamma C, & p^* &= -p; \\ C\gamma_4 &= -\gamma_4 C, & p_4^* &= (ip_0)^* = p_4 \end{aligned} \right\} \quad (11.3)$$

and the potentials of the external field are real,

$$A^* = A; \quad A_4^* = (A_0)^* = -A_4,$$

then introducing the positron function

$$\psi^{(p)} = C\psi^*,$$

we obtain

$$(i\gamma p + ie\gamma A + m)\psi^{(p)} = 0. \quad (11.4)$$

We see that Equation (11.4) differs from (11.1) in that the charge has the opposite sign. Thus, instead of considering solutions with positive and negative frequencies, we can consider only solutions with positive frequencies but for two signs of the charge. This is in agreement with our original concepts with respect to interpreting negative frequencies, as described in Section 8.

2. Separation of Variables in a Central Field.

Let us consider the problem of electron motion in a centrally symmetric electromagnetic field given by

$$A = 0; \quad A_0 = A_0(r).$$

The Hamiltonian (11.2) in this case has the form

$$H = \alpha p + \beta m + V(r), \quad (11.5)$$

where

$$V = eA_0.$$

In view of the spherical symmetry of the field, the angular momentum and the inversion operator commute with the Hamiltonian. Therefore, there exist simultaneous eigenstates of the energy, angular momentum, and parity. The corresponding wave functions, as in Section 10, shall be denoted by

$$\psi_{jIM} = \begin{pmatrix} \varphi_{jIM} \\ \chi_{jIM} \end{pmatrix},$$

and the factor e^{-iEt} , which gives the time dependence, shall be dropped.

The eigenfunctions of the angular momentum and of the inversion operator have been found in Section 10. Thus, we know the angular dependence of the wave function, namely

$$\left. \begin{aligned} \varphi_{jIM} &= a(r) \Omega_{jIM}, \\ \chi_{jIM} &= b(r) \Omega_{jIM} \quad (l' = 2j - l), \end{aligned} \right\} \quad (11.6)$$

where Ω_{jIM} is a spherical spinor defined in (10.6). The problem reduces to the determination of the radial functions $a(r)$ and $b(r)$.

The equations satisfied by $a(r)$ and $b(r)$ will be obtained if we insert (11.6) into the eigenvalue equation of the energy operator

$$H\psi = e\psi,$$

i.e., into the set of equations

$$\left. \begin{aligned} (e - m - V)\varphi &= \sigma p \chi, \\ (e + m - V)\chi &= \sigma p \varphi. \end{aligned} \right\} \quad (11.7)$$

Let us transform the right sides of (11.7). In the second of these there occurs the expression

$$\sigma p \varphi = \sigma p a(r) \Omega_{jIM} \left(\frac{r}{r} \right) = -i \frac{da}{dr} \left(\sigma \frac{r}{r} \right) \Omega_{jIM} + \sigma p \Omega_{jIM}.$$

The left side of this expression contains the function χ , whose angular dependence is given by

$$\Omega_{JM}.$$

Using Equation (10.13) which gives the relation between Ω_{JM} and $\Omega_{JM'}$, we obtain

$$\begin{aligned} \left(\sigma \frac{r}{r}\right) \Omega_{JM} &= \Omega_{JM'}, \\ (\sigma p) \Omega_{JM} &= (\sigma p) \left(\sigma \frac{r}{r}\right) \Omega_{JM'} = \\ &= \left(p \frac{r}{r} + i \left(p \frac{r}{r} \sigma\right)\right) \Omega_{JM'} = -i \left(\frac{2}{r} + \frac{1}{r} L \sigma\right) \Omega_{JM'}, \end{aligned}$$

where

$$L = [r p]$$

is the orbital angular momentum operator. Since Ω_{JM} is an eigenfunction of the operators L^2 , $M^2 = (L + \frac{\sigma}{2})^2$ and $(\frac{\sigma}{2})^2$ belonging to the eigenvalues $l(l+1)$, $j(j+1)$, and $1/2(1/2+1)$, we arrive at

$$L \sigma \Omega_{JM} = \left(M^2 - L^2 - \left(\frac{\sigma}{2}\right)^2\right) \Omega_{JM} = \left\{j(j+1) - l(l+1) - \frac{3}{4}\right\} \Omega_{JM}.$$

Here $l = 2j - 1$, and

$$j(j+1) - l(l+1) - \frac{3}{4} = \begin{cases} -l+2 & \text{for } j = l + \frac{1}{2}, \\ l-1 & \text{for } j = l - \frac{1}{2}. \end{cases}$$

Let us introduce the following notation

$$x = \mp \left(j + \frac{1}{2}\right) = \begin{cases} -l-1 & \text{for } j = l + \frac{1}{2}, \\ l & \text{for } j = l - \frac{1}{2}. \end{cases} \quad (11.8)$$

(When $l = 0$ only the one value $j = 1/2$ is possible and, therefore, $x = -1$.) Then

$$(2 + L \sigma) \Omega_{JM} = (1 + x) \Omega_{JM} \quad \left(j = l \pm \frac{1}{2}\right)$$

and

$$\sigma p \varphi = -i \Omega_{JM} \left(\frac{da}{dr} + \frac{x+1}{r} a\right).$$

Similarly, for the right side of the first of Equations (11.7) we obtain

$$\sigma p \chi = -i \Omega_{JM} \left(\frac{db}{dr} - \frac{x-1}{r} b\right).$$

Inserting these expressions into (11.7) and cancelling the angular functions on both sides, we obtain the desired equations for the radial functions. In order to make these equations real, let us introduce the notation

$$\begin{cases} a(r) = ig(r), \\ b(r) = f(r). \end{cases} \quad (11.9)$$

Then f and g will satisfy

$$\begin{cases} \frac{dg}{dr} + (1+x) \frac{g}{r} - (e+m-V)f = 0, \\ \frac{df}{dr} + (1-x) \frac{f}{r} + (e-m-V)g = 0. \end{cases} \quad (11.10)$$

Equations (11.10) can be rewritten in the form

$$\begin{cases} \frac{d(rg)}{dr} + \frac{x}{r}(rg) - (e+m-V)(rf) = 0, \\ \frac{d(rf)}{dr} - \frac{x}{r}(rf) + (e-m-V)(rg) = 0. \end{cases} \quad (11.11)$$

3. Asymptotic Behavior.

In addition to Equations (11.11), rg and rf must also satisfy two boundary conditions which make it possible to normalize the wave function:

$$\left. \begin{aligned} \frac{rf}{rg} &\rightarrow 0 \quad \text{as } r \rightarrow 0, \\ \frac{rf}{rg} &\neq \infty \quad \text{as } r \rightarrow \infty. \end{aligned} \right\} \quad (11.12)$$

Let us find the asymptotic behavior of g and f . Assuming that

$$V \rightarrow 0 \quad \text{as } r \rightarrow \infty,$$

we obtain, as $r \rightarrow \infty$,

$$\begin{aligned} \frac{d}{dr}(rg) - (e+m)rf &= 0, \\ \frac{d}{dr}(rf) + (e-m)rg &= 0. \end{aligned}$$

The general solution of these equations is

$$\left. \begin{aligned} rg &= C(e^{-\lambda r} + \eta e^{\lambda r}), \\ rf &= C\left(-\sqrt{\frac{m-e}{m+e}}e^{-\lambda r} + \eta\sqrt{\frac{m-e}{m+e}}e^{\lambda r}\right), \end{aligned} \right\} \quad (11.13)$$

where C and η are constants (strictly speaking C is a slowly varying function of r), and

$$\lambda = \sqrt{m^2 - e^2}. \quad (11.14)$$

Neither constant in (11.13) is arbitrary. The constant C is determined by the normalization of the wave function. The constant η is determined by the fact that (11.13) should be the asymptotic expression for those solutions of (11.11) which satisfy the first boundary condition (11.12). We find that it is a function of the energy

$$\eta = \eta(e),$$

and also of the angular momentum.

The behavior of the functions given by (11.13) is quite different for $e > m$ and $e < m$. When $e > m$ the coefficient in the exponent of (11.13) is seen by (11.14) to be imaginary. In this case the second of the boundary conditions (11.12) is satisfied independent of the value of e . The asymptotic expression for the radial functions (11.13) can then be written, as is easy to see, in the form

$$\left. \begin{aligned} rg &= c_1 \sqrt{1 + \frac{e}{m}} \cos(pr + \delta'), \\ rf &= c_1 \sqrt{1 - \frac{e}{m}} \sin(pr + \delta'), \end{aligned} \right\} \quad (11.15)$$

where

$$p = -i\lambda = \sqrt{e^2 - m^2}, \quad (11.16)$$

and c_1 is a constant determined by the normalization. The phase angle δ' depends on the external field as well as on the energy and angular momentum of the electron.

When $e < m$, the quantity λ is real, and therefore, the second terms in (11.13) increase exponentially. Then the solutions given by (11.13) satisfy the boundary condition at infinity only if

$$\eta(e) = 0. \quad (11.17)$$

The roots of (11.17) determine the possible energy values. Thus, as has been previously asserted, we obtain a continuous spectrum when $e > m$, and a discrete spectrum when $e < m$. The latter may be absent if (11.17) has no roots.

4. Level Behavior as a Function of the Potential-Well Depth.

Let us now return to a consideration of the possibility of separating the solutions of the Dirac equation in an external field into electron and positron states. For this purpose let us consider the frequency spectrum in the special case of a spherically symmetric "potential well".¹⁾

$$\begin{aligned} V &= -V_0 & \text{for } r < r_0, \\ V &= 0 & \text{for } r > r_0, \end{aligned}$$

where V_0 is a constant. This simple example can be used to clarify several general aspects of the discrete spectrum. Equations (11.1) become

¹⁾ L. Schiff, H. Snyder, and J. Weinberg, Phys. Rev. 57, 315 (1940).

$$\left. \begin{aligned} \frac{d}{dr}(rg) + \frac{\omega}{r}(rg) - (\omega + m + V_0)(rf) &= 0, \\ \frac{d}{dr}(rf) - \frac{\omega}{r}(rf) + (\omega - m + V_0)(rg) &= 0; \\ \frac{d}{dr}(rg) + \frac{\omega}{r}(rg) - (\omega + m)(rf) &= 0, \\ \frac{d}{dr}(rf) - \frac{\omega}{r}(rf) + (\omega - m)(rg) &= 0; \end{aligned} \right\} \begin{aligned} r < r_0 \\ r > r_0 \end{aligned} \quad (11.18)$$

The first of the boundary conditions (11.12) now refers to the first two of these equations (the region $r < r_0$), and the second one refers to the second two equations. In addition, f and g should satisfy the continuity condition at $r = r_0$. Instead of the second boundary condition of (11.12):

$$r f, r g \neq \infty \quad \text{as} \quad r \rightarrow \infty$$

we shall use the condition

$$f = g = 0 \quad \text{for} \quad r = R \quad (R \gg r_0).$$

This corresponds to placing the system into a large spherical "box" with impenetrable walls. In principle, changing the boundary condition in this way has no particular meaning. It is, however, convenient in practice because it makes the whole spectrum discrete (with very small separation between adjoining frequencies in the region $|\omega| > m$). In this way we can follow the variation of each level as we vary the depth of the potential well V_0 .

We shall restrict ourselves to the case $\kappa = -1$. Eliminating f from Equations (11.18), we obtain a second order equation for g , namely

$$\left. \begin{aligned} \frac{d^2}{dr^2}(rg) + [(\omega + V_0)^2 - m^2](rg) &= 0, & r < r_0 \\ \frac{d^2}{dr^2}(rg) + (\omega^2 - m^2)(rg) &= 0, & r > r_0 \end{aligned} \right\}$$

Those solutions of these equations which satisfy the boundary conditions at $r = 0$ and $r = R$ are

$$\left. \begin{aligned} rg &= c_1 \sin(\sqrt{(\omega + V_0)^2 - m^2} r), & r < r_0 \\ rg &= c_2 \sin(\sqrt{\omega^2 - m^2} (R - r)), & r > r_0 \end{aligned} \right\}$$

where c_1 and c_2 are constants. We note that the argument of the sine is real or imaginary depending on whether or not the following expressions are positive or negative:

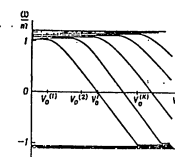


Fig. 1

The function f is determined by g from (11.18) in both regions. It then also satisfies the boundary conditions at $r = r_0$ and at $r = R$. Thus, the condition at the boundary $r = r_0$, when written in the form

$$\left(\frac{f}{g}\right)_{r=r_0-0} = \left(\frac{f}{g}\right)_{r=r_0+0}, \quad (11.19)$$

contains no arbitrary constants. It is the characteristic equation which determines the possible values of ω and replaces (11.17) in our case.

Figure 1 is a schematic diagram of the dependence of the frequency spectrum on the potential well depth V_0 for a fixed potential well radius r_0 , as obtained from the solutions of Equation (11.19). We see from the graph that when $V_0 < V_0^m$ (1), no bound states ($|\omega| < m$) exist. The frequency spectrum consists, as in the absence of an external field, of two regions $\omega > m$ and $\omega < m$ (we shall call them the upper and lower continua). When $V_0 > V_0^m$ (2), the lowest level of the upper continuum lies below m , which means that there exists a single bound state; when $V_0 = V_0^m$ (3), there appears a second bound state, etc. The frequency of each bound state decreases continuously as V_0 increases, becoming negative at $V_0 = V_0^m$. Nevertheless, these states can be considered as electron states, since by adiabatically varying the external field we can "return" this state to the upper continuum. A difficulty arises at $V_0 = V_0^m$ (4), when the level crosses the boundary $\omega = -m$ and joins the lower continuum, which is the set of positron states. The problem of the behavior of an electron in a potential well whose depth is greater than V_0^m (5) cannot be solved within the framework of single-particle quantum mechanics. Later (see Section 17) we shall show that in this case the external field absorbs the electron.

§ 12. Electron Motion in the Field of the Nucleus

1. Solution of the Radial Equation for the Coulomb Field. Discrete Spectrum.

A most important application of the Dirac equation is the study of the motion of the electron in the field of the nucleus. The latter is not strictly a central field. The deviation from spherical symmetry is due to the fact that nuclei generally possess electric quadrupole and magnetic dipole (as well as higher multipole) moments. If we neglect these effects (which produce a hyperfine structure in the electron levels), then we can consider electron states which are simultaneously eigenstates of the energy, angular momentum, and parity. The wave function then is of the form (11.6), and the problem reduces to the solution of the equations for the radial functions (11.11).

At large distances the nuclear field is Coulombic¹⁾

$$A_0 = -\frac{Ze}{r}, \quad V = -\frac{Ze^2}{r} \quad (r \gg r_0), \quad (12.1)$$

¹⁾ In this section and the two following, we are using Gaussian units, rather than Heaviside.

where Z_0 is the nuclear charge, and r_0 is its radius. We shall first consider the problem in the approximation in which the finite nuclear dimensions may be neglected, and shall consider expression (12.1) valid in all space.

In agreement with the asymptotic behavior (11.13) of the radial functions, we shall attempt to find solutions of Equations (11.11) in the form

$$\left. \begin{aligned} rG &= \sqrt{1 + \frac{1}{m}} e^{-\lambda r} (F_1 + F_2), \\ rF &= \sqrt{1 - \frac{1}{m}} e^{-\lambda r} (F_1 - F_2). \end{aligned} \right\} \quad (12.2)$$

Inserting (12.2) and (12.1) into (11.11), and introducing the independent variable

$$\rho = 2\lambda r \quad (12.3)$$

we arrive at the following equations for F_1 and F_2 :

$$\left. \begin{aligned} \frac{dF_1}{d\rho} &= \left(1 - \frac{1}{\lambda} \frac{Ze^2}{\rho}\right) F_1 - \left(\frac{1}{\rho} + \frac{m}{\lambda} \frac{Ze^2}{\rho}\right) F_2, \\ \frac{dF_2}{d\rho} &= \left(-\frac{1}{\rho} + \frac{m}{\lambda} \frac{Ze^2}{\rho}\right) F_1 + \left(\frac{1}{\rho} + \frac{1}{\lambda} \frac{Ze^2}{\rho}\right) F_2. \end{aligned} \right\} \quad (12.4)$$

The solution of (12.4) for small ρ is of the form

$$\left. \begin{aligned} F_1 &= a_1 \rho^b, \\ F_2 &= a_2 \rho^b. \end{aligned} \right\} \quad (12.5)$$

Inserting (12.5) into (12.4) we obtain algebraic equations for a_1 and a_2 , namely

$$\left. \begin{aligned} a_1 \left(\gamma + Ze^2 \frac{m}{\lambda} \right) + a_2 \left(\gamma + Ze^2 \frac{m}{\lambda} \right) &= 0, \\ a_1 \left(\gamma - Ze^2 \frac{m}{\lambda} \right) + a_2 \left(\gamma - Ze^2 \frac{m}{\lambda} \right) &= 0. \end{aligned} \right\}$$

Setting the determinant of this set of equations equal to zero, we arrive at

$$\gamma = \pm \sqrt{x^2 - Z^2 e^4}. \quad (12.6)$$

Only the + sign in (12.6) is compatible with the first boundary condition of (11.12). We shall, therefore, henceforth use the positive value of γ .

Thus, $\rho^{-\gamma} F_1$ and $\rho^{-\gamma} F_2$ are finite at $\rho = 0$. If we eliminate one of the functions, say F_1 , from (12.4), we obtain a second order equation for $\rho^{-\gamma} F_2$, whose form is

$$\left[\rho \frac{d^2}{d\rho^2} + (b - \rho) \frac{d}{d\rho} - a \right] (\rho^{-\gamma} F_2) = 0, \quad (12.7)$$

where

$$b = 2\gamma + 1; \quad a = \gamma - Ze^2 \frac{m}{\lambda}.$$

The solution of (12.7) which is finite for $\rho = 0$ is, as is well known, the confluent hypergeometric function, which can be represented by the series

$$F(a, b; \rho) = \frac{\Gamma(b)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)}{\Gamma(b+n)} \frac{\rho^n}{n!}.$$

Thus,

$$F_2 = c (2\lambda r)^\gamma F\left(\gamma - Ze^2 \frac{m}{\lambda}, 2\gamma + 1; 2\lambda r\right), \quad (12.8)$$

where c is a constant. Further, according to (12.4) F_1 can be expressed in terms of F_2 and $\frac{dF_2}{d\rho}$. Using a recursion formula for the hypergeometric function

$$\rho \frac{d}{d\rho} F(a, b; \rho) = a \{ F(a+1, b; \rho) - F(a, b; \rho) \},$$

we obtain

$$F_1 = c \frac{\gamma - Ze^2 \frac{m}{\lambda}}{-x + Ze^2 \frac{m}{\lambda}} (2\lambda r)^{\gamma} F(\gamma+1, -Ze^2 \frac{m}{\lambda}, 2\gamma+1; 2\lambda r). \quad (12.9)$$

We have seen that when $\epsilon < \infty$, that is for real values of λ , these solutions have meaning only for those discrete values of ϵ which are given by (11.17). In order to find these values, we shall make use of the asymptotic expression for the hypergeometric function

$$F(a, b; p) \approx e^{-\epsilon a} \frac{\Gamma(b)}{\Gamma(b-a)} p^{-a} + \frac{\Gamma(b)}{\Gamma(a)} p^{a-b} e^{\epsilon p}. \quad (12.10)$$

It follows from this that in order for the exponentially increasing terms in F_1 and F_2 to vanish,

$$\left. \begin{aligned} \frac{1}{\Gamma(\gamma - Ze^2 \frac{m}{\lambda})} &= 0, \\ \frac{\gamma - Ze^2 \frac{m}{\lambda}}{-x + Ze^2 \frac{m}{\lambda}} \frac{1}{\Gamma(\gamma - Ze^2 \frac{m}{\lambda} + 1)} &= \frac{1}{(Ze^2 \frac{m}{\lambda} - x) \Gamma(\gamma - Ze^2 \frac{m}{\lambda})} = 0. \end{aligned} \right\} \quad (12.11)$$

If $Ze^2 \frac{m}{\lambda} - x \neq 0$, then for $\kappa < 0$, the second of conditions (12.11) is identical with the first. Since the poles of the Γ -function are the negative integers and zero, (12.11) becomes

$$\gamma - Ze^2 \frac{m}{\lambda} = -n', \quad (12.12)$$

where n' is a nonnegative integer. When $x = Ze^2 \frac{m}{\lambda}$ (12.6) gives

$$\gamma = Ze^2 \frac{m}{\lambda},$$

which means that $n' = 0$. Then the second of conditions (12.11) is not satisfied. Thus,

$$\left. \begin{aligned} n' &= 0, 1, 2, \dots & \text{when } x < 0, \\ n' &= 1, 2, \dots & \text{when } x > 0. \end{aligned} \right\} \quad (12.13)$$

Solving (12.12) for ϵ , we obtain

$$\epsilon = \frac{m}{\sqrt{1 + (n' + \gamma)^2}}. \quad (12.14)$$

Equation (12.14) gives the fine structure of the hydrogen atom. When $Ze^2 \ll 1$, an expansion in this parameter up to $(Ze^2)^4$ leads to

$$\frac{\epsilon - m}{m} = -\frac{(Ze^2)^2}{2n^2} - \frac{(Ze^2)^4}{2n^3} \left(\frac{1}{|x|} - \frac{3}{4n} \right), \quad (12.15)$$

where³⁾

$$n = n' + |x|.$$

The number n is the "principal quantum number" of the nonrelativistic theory; the first term of (12.15) gives the Balmer series.

An interesting property of the fine structure formula (12.14) [and, therefore, also of (12.15)] is the fact that only the absolute value of κ enters into the expression [it enters in terms of γ by Equation (12.6)]. Thus, all the energy levels except the lowest, for which according to (12.13)

$$n' = 0, \quad x = -1,$$

are doubly degenerate. In particular, the following levels are identical:

$$2s_{1/2} (n' = 1, \quad x = -1; \quad l = 0, \quad j = \frac{1}{2}; \quad n = 2)$$

³⁾ We note that $a_0 = 1/m\epsilon^2$ is the Bohr radius of the lowest electron state in the hydrogen atom. Equation (12.15) can also be written in the form

$$\frac{\epsilon - m}{m} = -\frac{Ze^2}{2a_0 n^2} \left[1 - \frac{(Ze^2)^2}{n} \left(\frac{1}{|x|} - \frac{3}{4n} \right) \right].$$

and

$$2p_{10} \left(n' = 1, \quad x = 1; \quad l = 1, \quad j = \frac{1}{2}, \quad n = 2 \right).$$

We shall see later (see Section 46) that Equation (12.14), which is based only on the concept of a given external field, is not exact. Corrections to this formula partially remove the degeneracy in κ . These corrections for small values of Z are of order of magnitude $(Ze^2)^2 \frac{1}{a_0^3}$. It is thus seen that in hydrogen, for instance, the use of Equation (12.14) with an accuracy greater than that given in (12.15) is not necessary.

In order to completely determine the wave function of the discrete spectrum we must still find the value of the constant c in (12.8). This is found from the normalization condition

$$\int_0^\infty (g^2 + f^2) r^2 dr = 1.$$

Carrying out the integration we obtain¹⁾

$$c = \sqrt{\frac{Ze^2 \frac{m}{\lambda} - x}{2Ze^2 \frac{m}{\lambda}} \frac{\Gamma(2\gamma + n' + 1)}{\Gamma(2\gamma + 1) \sqrt{\pi!}} \sqrt{\lambda}}. \quad (12.16)$$

We see that the radial functions depend on the two quantum numbers n' and κ , and that they are given by

$$g = \frac{\sqrt{\Gamma(2\gamma + n' + 1)}}{\Gamma(2\gamma + 1) \sqrt{\pi!}} \sqrt{\frac{1 + \frac{e}{m}}{4N(N-x)}} \left(\frac{2Z}{Na_0} \right)^{1/2} e^{-\frac{Zr}{Na_0}} \left(\frac{2Zr}{Na_0} \right)^{1-\gamma} \times \\ \times \left\{ n' F(-n' + 1, 2\gamma + 1; \frac{2Zr}{Na_0}) - (N-x) F(-n', 2\gamma + 1; \frac{2Zr}{Na_0}) \right\}; \quad (12.17)$$

$$f = -\frac{\sqrt{\Gamma(2\gamma + n' + 1)}}{\Gamma(2\gamma + 1) \sqrt{\pi!}} \sqrt{\frac{1 - \frac{e}{m}}{4N(N-x)}} \left(\frac{2Z}{Na_0} \right)^{1/2} e^{-\frac{Zr}{Na_0}} \left(\frac{2Zr}{Na_0} \right)^{1-\gamma} \times \\ \times \left\{ n' F(-n' + 1, 2\gamma + 1; \frac{2Zr}{Na_0}) + (N-x) F(-n', 2\gamma + 1; \frac{2Zr}{Na_0}) \right\}, \quad (12.17')$$

¹⁾ See H. Bethe, Quantum Mechanics of Simple Systems (United Sci. Tech. Press, 1935) [This is a translation of "Quantenmechanik der Ein und Zwei Elektronenprobleme," Handbuch der Physik, 2nd Ed., XXIV, Part 1 (1933)].

where

$$N = \sqrt{n^2 - 2n'(1-x-\gamma)}, \quad a_0 = \frac{1}{me^2}, \quad n = n' + |x|; \\ \gamma = \sqrt{x^2 - Z^2 a_0^2}, \quad \alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}.$$

2. Wave Functions of the Continuous Spectrum.

Let us go on to a consideration of the continuous spectrum, with $\epsilon > \underline{m}$. Then λ is pure imaginary,

$$\lambda = ip, \quad (12.18)$$

where $p = \sqrt{\epsilon^2 - m^2}$.

Equations (12.8) and (12.9) remain valid. We shall, however, write them in a somewhat different form. First, we note that the factor

$$\frac{\gamma - Ze^2 \frac{m}{\lambda}}{-x + Ze^2 \frac{m}{\lambda}} = -\frac{\gamma + iZe^2 \frac{m}{p}}{x + iZe^2 \frac{m}{p}}$$

in (12.9) now has modulus unity. It is, therefore, convenient to introduce the notation

$$e^{2i\eta} = -\frac{x - iZe^2 \frac{m}{p}}{\gamma + iZe^2 \frac{m}{p}}, \quad (12.19)$$

where η is real. Including the common factor in the constant, we can write (12.8) and (12.9) in the form

$$F_1 = Ce^{i\eta} (\gamma + i\nu) F(\gamma + 1 + i\nu, 2\gamma + 1; 2ipr) (2pr)^\gamma, \\ F_2 = Ce^{-i\eta} (\gamma - i\nu) F(\gamma + i\nu, 2\gamma + 1; 2ipr) (2pr)^\gamma, \quad (12.20)$$

where

$$\nu = Ze^2 \frac{e}{p} = \frac{Ze^2}{v}$$

(v is the velocity of the electron).

The two hypergeometric functions in (12.20) are simply related to each other. This relation may be obtained by using the integral representation of the hypergeometric function, namely

$$F(a, b; x) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_{-1}^1 (1+z)^{a-1} (1-z)^{b-a-1} e^{\frac{x}{2}(1+z)} dz. \quad (12.21)$$

If x is imaginary, and the real part of a is related to b by

$$2\operatorname{Re} a + 1 = b,$$

then it can be shown from (12.21) that

$$F(a, b; x) = e^{\pi i} [F(a+1, b; x)]^*.$$

Making use of this equation and inserting the expressions for F_1 and F_2 into (12.2), we arrive at the radial functions of the form

$$\left. \begin{aligned} rg &= c \sqrt{1 + \frac{1}{m}} (2pr)^{\frac{1}{2}} \times \\ &\times [e^{-ipr+iv} (\gamma+iv) F(\gamma+1+iv, 2\gamma+1; 2ipr) + \\ &+ e^{ipr-iv} (\gamma-iv) F^*(\gamma+1+iv, 2\gamma+1; 2ipr)]; \\ rf &= ci \sqrt{-1 + \frac{1}{m}} (2pr)^{\frac{1}{2}} \times \\ &\times [e^{-ipr+iv} (\gamma+iv) F(\gamma+1+iv, 2\gamma+1; 2ipr) - \\ &- e^{ipr-iv} (\gamma-iv) F^*(\gamma+1+iv, 2\gamma+1; 2ipr)]. \end{aligned} \right\} \quad (12.22)$$

From (12.22) it can be seen that g and f are real if the constant c is chosen real.

Let us now find the asymptotic expressions for the functions in (12.22). For this purpose let us use Equation (12.10). Since now $\rho = 2ipr$ is pure imaginary, the decrease of the hypergeometric function is determined not by the exponential but by the power factors in (12.10).

Since

$$\operatorname{Re} a = \gamma + 1, \quad b = 2\gamma + 1,$$

the first term decreases more rapidly than the second, and we may, therefore, consider only the second term. Thus,

$$\begin{aligned} F(\gamma+1+iv, 2\gamma+1; 2ipr) &\approx \frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+iv+1)\Gamma(2pr)} (2ipr)^{\gamma+iv} e^{2ipr} = \\ &= \frac{\Gamma(2\gamma+1)}{(\gamma+iv)\Gamma(\gamma+iv)\Gamma(2pr)} e^{-\frac{\pi}{2}} e^{i(\gamma pr - \frac{\pi}{2}\gamma + i\gamma pr)}. \end{aligned}$$

Inserting this expression in (12.22) we obtain

$$\left. \begin{aligned} rg &\approx c \sqrt{1 + \frac{1}{m}} \frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+iv)} e^{-\frac{\pi}{2}} e^{i\gamma pr} 2 \cos(pr + \delta'), \\ rf &\approx -c \sqrt{-1 + \frac{1}{m}} \frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+iv)} e^{-\frac{\pi}{2}} e^{i\gamma pr} 2 \sin(pr + \delta'), \end{aligned} \right\} \quad (12.23)$$

where

$$\delta' = \eta - \arg \Gamma(\gamma+iv) - \pi \frac{\gamma}{2} + \gamma \ln 2pr. \quad (12.24)$$

Equation (12.23) agrees with the general asymptotic expression (11.15) for the radial functions. The phase δ' is found not to be a constant, but a logarithmic function of r , which is a characteristic property of the Coulomb field and is related to its slow decrease with distance.

We shall find the constant c from the normalization condition written in the form

$$\int_0^\infty (g, g_{e'} + f, f_{e'}) r^2 dr = \delta(e - e').$$

Since the integrand does not go to zero as $r \rightarrow \infty$, we may insert the asymptotic expression (12.23) for the radial functions into this integral; the quadrature is elementary, and we obtain

$$c = \frac{e^{-\frac{\pi}{2}}}{\sqrt{\pi p}} \frac{|\Gamma(\gamma+iv)|}{2\Gamma(2\gamma+1)}. \quad (12.25)$$

The transition to other normalizations is accomplished just as for the free electron. When normalizing for an energy interval $d\epsilon$, (12.25) must be multiplied by $\sqrt{d\epsilon}$. When normalizing for a sphere of radius R (12.25) is multiplied by $\sqrt{\frac{\pi p}{4R}}$.

3. Isotope Level Shift.

The results thus far obtained are based on the application of the Coulomb law for all values of r . When the wave function of the electron is large in the neighborhood of the nucleus, we must take account of the fact that the potential given by (12.1) is valid only for $r > r_0$ (where r_0 is the nuclear radius).

An example of the effects due to finite nuclear dimensions is the isotope level shift, in which different electron energy levels are obtained for nuclei with the same charge. We shall show that the isotope shift can be expressed directly in terms of the value of the wave function at $r = r_0$.

Let ϵ be the energy of the electron in a Coulomb field, and $\epsilon + \epsilon_1$ be the energy in the field of a finite nucleus. We shall denote the radial wave functions in the first case by f_0 and g_0 , and in the second case by f and g . Let us consider the case $\kappa = -1$. According to (11.11) f_0 and g_0 satisfy the equations

$$\left. \begin{aligned} \frac{d}{dr}(rg_0) - \frac{1}{r}(rg_0) &= (\epsilon + m - V)(rf_0), \\ \frac{d}{dr}(rf_0) + \frac{1}{r}(rf_0) &= (m + V - \epsilon)(rg_0), \end{aligned} \right\} \quad (12.26)$$

where V is the Coulomb field (12.1). The functions f and g for $r > r_0$ satisfy the equations

$$\left. \begin{aligned} \frac{d}{dr}(rg) - \frac{1}{r}(rg) &= (\epsilon + m - V)(rf) + \epsilon_1 rf, \\ \frac{d}{dr}(rf) + \frac{1}{r}(rf) &= (m + V - \epsilon)(rg) - \epsilon_1 rg. \end{aligned} \right\} \quad (12.27)$$

Let us multiply the first of Equations (12.26) by rf , the second by $-rg$, the first of Equations (12.27) by $-rf_0$, and the second by rg_0 ; then adding all four equations we obtain

$$\frac{d}{dr}(gfg - fgf) = \epsilon_1 r^2 (g_0g + f_0f). \quad (12.28)$$

Let us integrate (12.28) over r from r_0 to ∞ . Assuming that f_0 and g_0 differ from f and g only in a small region close to the nucleus, and that this region gives a small contribution to the normalizing integral of the wave function, we can write approximately

$$\int_{r_0}^{\infty} r^2 (gfg + fgf) dr \approx \int_{r_0}^{\infty} (g_0^2 + f_0^2) r^2 dr = 1.$$

Thus,

$$\epsilon_1 = r_0^2 [g_0(r_0)f(r_0) - g(r_0)f_0(r_0)]. \quad (12.29)$$

The isotope shift is given by the difference of expressions such as (12.29) for two nuclear radii r_0 corresponding to two isotopes.

4. General Investigation of the Effect of Finite Nuclear Dimensions.

The difference between the actual wave function and the Coulomb wave function may be significant also for the continuous electron spectrum if the effect is due to the behavior of the wave function close to the nucleus. This situation arises, for instance, in the emission of β -electrons from the nucleus.¹⁾

The change in the wave function may be quite large. In order to evaluate it, let us return to the expressions for the radial functions (11.11) in the general case of an arbitrary central field.

The behavior for small r can be found in the following way.²⁾ Let us construct formal solutions of each of Equations (11.11), treating the last term as an inhomogeneity. This gives

$$\left. \begin{aligned} rg &= c_1 r^{-\kappa} + r^{-\kappa} \int_0^r r^{\kappa} (\epsilon + m - V) rf dr, \\ rf &= c_2 r^{\kappa} - r^{\kappa} \int_0^r r^{-\kappa} (\epsilon - m - V) rg dr. \end{aligned} \right\} \quad (12.30)$$

From the boundary condition at $r = 0$ we obtain

$$\left. \begin{aligned} c_1 &= 0 & \text{for } \kappa > 0, \\ c_2 &= 0 & \text{for } \kappa < 0. \end{aligned} \right\}$$

Thus, for small r we have

$$\left. \begin{aligned} rg &= c_1 r^{1+\kappa}, \\ rf &= -c_1 r^{-1+\kappa} \int_0^r r^{2+\kappa} (\epsilon - m - V) dr; & \kappa < 0, \\ rf &= c_2 r^{\kappa}, \\ rg &= c_2 r^{-\kappa} \int_0^r r^{2-\kappa} (\epsilon + m - V) dr; & \kappa > 0. \end{aligned} \right\} \quad (12.30')$$

¹⁾ Ya. Smorodinsky, J. Exptl.-Theoret. Phys. 17, 1034 (1947).

²⁾ L. Siliv, J. Exptl.-Theoret. Phys. 17, 1049 (1947).

³⁾ H. Rose, Phys. Rev. 82, 389 (1951).

We see that when $\kappa < 0$ the dependence of g on r is determined by κ , or the angular momentum of the electron, and that in the first approximation it is independent of the field. The function f , on the other hand, depends largely on the field and thus, the difference between the actual field and the Coulomb field effects its behavior. Since for small r (for nuclei with not very small Z)

$$|V| \gg m,$$

we obtain, for electron energies that are not too high,

$$rf = c_1 r^{-1} \int_0^r r'^{2(1+\kappa)} V dr', \quad \kappa < 0.$$

When $\kappa > 0$ the situation is similar, except that the roles of f and g are interchanged.

In order that a separation of the form (12.30') have meaning, it is necessary that the integral obtained when (12.30') is substituted into (12.30) exists. This will be true for all κ , so long as $\lim_{r \rightarrow 0} rV = 0$. Thus, the separation (12.30) is not applicable to a pure Coulomb field. However, the field of a nucleus of finite dimensions remains finite (or goes to zero) when $r = 0$, and therefore, (12.30) may be used.

We shall not give a complete treatment of the equations for the radial functions about a nucleus of finite dimensions. We shall indicate only the basic properties of this problem.

First of all it should be noted that the function $V(r)$ is not well defined for $r < r_0$, and the boundary of the nucleus is to some extent merely conventional. A qualitatively correct description will be obtained if we assume that the charge density is constant within a sphere of radius r_0 . Then

$$V = -\left(\frac{3}{2} - \frac{1}{2} \frac{r^2}{r_0^2}\right) \frac{Ze^2}{r_0}, \quad r < r_0. \quad (12.31)$$

Another limiting case is that when the charge is distributed on the surface of a sphere of radius r_0 . In this case

$$V = V_0 = -\frac{Ze^2}{r_0}, \quad r < r_0. \quad (12.32)$$

The fields given by (12.31) and (12.32) lead to similar results for the change in the wave function in the region of r_0 when $r > r_0$. What is important here is not the actual variation of the potential for $r < r_0$, but the fact that its Coulomb increase ceases as $r \rightarrow 0$.

The radial functions corresponding to the case (12.32) (constant potential) are the same as those for the free electron, except that ϵ is replaced by $\epsilon + V_0$. In practice it is sufficient to use the first terms of the series expansions for f and g in terms of r . For this we can make use of Equation (12.30) for both cases (12.32) and (12.31). This determines the solution for $r < r_0$ up to a constant factor.

In the region $r > r_0$ the expression (12.1) for the potential is valid. Therefore, the solutions are of the form (12.2) with Equation (12.4). Since in this case, however, the point $r = 0$ is not included in the region of solution of (12.4), we may drop the boundary conditions which were satisfied in solutions (12.30) in the region $r < r_0$. Thus, we must maintain both signs in front of the radical in (12.6). We shall again use the notation

$$\gamma = +\sqrt{x^2 - z^2 a^2},$$

so that

$$F_2 = F_2(\gamma) + \zeta F_2(-\gamma),$$

where $F_2(\gamma)$ is given by (12.8), and ζ is a constant. Similarly,

$$F_1 = F_1(\gamma) + \zeta F_1(-\gamma),$$

where $F_1(\gamma)$ is given by (12.9).

The functions f and g thus depend on the two constants c and ζ , the first of these being a common factor. Thus, the ratio f/g contains only the constant ζ , which can be found from the continuity condition at $r = r_0$, namely

$$\left(\frac{f}{g}\right)_{r=r_0+0} = \left(\frac{f}{g}\right)_{r=r_0-0}.$$

We note that according to (12.30) the ratio f/g contains no undetermined constants when $r < r_0$.

The constant c is found from the normalization conditions. According to the general conclusions made in the previous section, when $\epsilon < m$ the boundary condition at $r = \infty$ gives a discrete spectrum.

We present the result of inserting the values of $f(r_0)$ and $g(r_0)$ so obtained for the potential given by (12.31) into the expression for the isotope shift (12.29). If $\Delta\epsilon$ is the level shift (for $\kappa = -1$) of the electron in the field of two nuclear isotopes whose radii differ by Δr_0 , then¹⁾

$$\Delta\epsilon = 8me^4 \frac{(n-1)!}{(n!)^2 \pi^4} \frac{\gamma^2}{[\Gamma(2\gamma+1)]^2} \frac{2-\gamma}{2+\gamma} (2Ze^2 m r_0)^{2\gamma} \frac{\Delta r_0}{r_0}. \quad (12.33)$$

5. On the Existence of Bound States for Large Z .

In concluding this section let us note that the lowest electron energy level ($\kappa = -1$, $n' = 0$) as given by (12.14) for $Ze^2 = 1$ ($Z = 137$) is $\epsilon = 0$.

¹⁾ Breit and Rosenzweig, Phys. Rev. 41, 451 (1932); Racah, Nature 129, 723 (1932).

When $Z\alpha^2 > 1$, Equation (12.14) gives imaginary values of ϵ , which shows that these states do not exist. However, this conclusion is valid only for the Coulomb field, and it is therefore incorrect to apply it to the field of the nucleus. Taking account of finite nuclear dimensions leads to the conclusion that bound states may exist also when $Z\alpha^2 > 1$. As Z increases $\epsilon_{\min}$ decreases and, as was true in the example of a simple potential well given at the end of the previous section, at some value $Z = Z(k)$ it becomes less than $-2m$. This value $Z(k)$ is

$$\begin{aligned} Z^{(k)} &= 200 & \text{for } r_0 &= 1.2 \cdot 10^{-12}, \\ Z^{(k)} &= 175 & \text{for } r_0 &= 8 \cdot 10^{-13}. \end{aligned}$$

Thus, according to the interpretation given at the end of Section 11, an electron in the field of the nucleus for which $Z > Z(k)$, would be captured by the nucleus. Then the nuclear charge would decrease by unity. This phenomenon would give the upper bound of the periodic system of elements, were it not true that this system ends at lower values for reasons related to the instability of heavy nuclei.

§ 13. Scattering of Electrons.

1. Born Approximation for Scattering in a Coulomb Field.

The problem of the scattering of particles in an external field can be solved by the same methods in relativistic as in nonrelativistic quantum mechanics.¹⁾ Let us start with the simplest case, in which the external field can be treated as a perturbation. In this case the Born approximation may be applied, and we write

$$d\sigma = \frac{2\pi}{v} \left| \int \psi_2^* V \psi_1 d\mathbf{r} \right|^2 \delta(\epsilon_2 - \epsilon_1) \frac{d\mathbf{p}}{(2\pi)^3}, \quad (13.1)$$

where $d\sigma$ is the differential scattering cross section, v is the velocity of the electron, and ψ_1 and ψ_2 are the free electron wave functions before and after scattering, corresponding to momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ and polarizations μ_1 and μ_2 , namely

$$\begin{aligned} \psi_1 &= u_1 e^{i\mathbf{p}_1 \cdot \mathbf{r}}, \\ \psi_2 &= u_2 e^{i\mathbf{p}_2 \cdot \mathbf{r}} \end{aligned}$$

(u_1 and u_2 are unit bispinors; see Section 9). If the external field has no vector potential, then

$$V = eA_0$$

and

¹⁾ I. Pomeranchuk and Ya. Smorodinsky, J. Phys. (USSR) 9, 97 (1945).

²⁾ See, for instance, D. Blokhintsev, Elements of Quantum Mechanics (State Tech. Press, 1949); L. Landau and E. Lifshitz, Quantum Mechanics (State Tech. Press, 1949); Mott and Massey, Theory of Atomic Collisions (Foreign Lit. Press, 1951).

$$d\sigma = \frac{1}{(2\pi)^3} \frac{p^2}{v^2} \left| \int \psi_2^* V \psi_1 d\mathbf{r} \right|^2 \delta(\epsilon_2 - \epsilon_1) \quad (13.2)$$

where $d\Omega$ is the element of solid angle in the direction in which the electron is scattered.

Let us average expression (13.2) over the initial spins, and sum over the final spins. This will give the differential cross section for scattering of unpolarized electrons. For this purpose, $|\psi_2^* \psi_1|^2$ in (13.2) should be replaced by $\frac{1}{4} \sum |\psi_2^* \psi_1|^2$, where $\sum$ denotes summation over initial and final spins. The method of summing such expressions was described in Section 9. According to (9.20), (9.23), and (9.24),

$$\begin{aligned} \sum |\psi_2^* \psi_1|^2 &= \sum |\bar{u}_2 \gamma_4 u_1|^2 = \\ &= \frac{1}{4\pi} \text{Sp} \gamma_4 (m - i\gamma p_1) \gamma_4 (m - i\gamma p_2) = \\ &= \frac{1}{4\pi} \text{Sp} (m^2 - p_1 p_2 \gamma_4 \gamma_4) = 2 \left(1 - v^2 \sin^2 \frac{\theta}{2} \right). \end{aligned}$$

Here $(\mathbf{p}_1 \mathbf{p}_2) = -\epsilon_1 \epsilon_2 + \mathbf{p}_1 \cdot \mathbf{p}_2 \cos \theta$, where θ is the scattering angle, $\epsilon_1 = \epsilon_2 = \epsilon$, and $|\mathbf{p}_1| = |\mathbf{p}_2| = p$. Thus,

$$d\sigma = \frac{1}{(2\pi)^3} \frac{p^2}{v^2} \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \left| \int e^{i(\mathbf{p}_1 - \mathbf{p}_2) \cdot \mathbf{r}} V d\mathbf{r} \right|^2 d\Omega. \quad (13.3)$$

Equation (13.3) differs from the corresponding nonrelativistic expression for the cross section (in addition to the relativistic dependence of p on v) by the factor

$$\left(1 - v^2 \sin^2 \frac{\theta}{2} \right).$$

In the case of the Coulomb field,

$$\int e^{i(\mathbf{p}_1 - \mathbf{p}_2) \cdot \mathbf{r}} V d\mathbf{r} = -\frac{4\pi Z e^2}{|\mathbf{p}_1 - \mathbf{p}_2|^2} = -\frac{4\pi Z e^2}{(2p \sin \frac{\theta}{2})^2}$$

and

$$d\sigma = \left(\frac{Z e^2}{2pv} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}} \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) d\Omega. \quad (13.4)$$

2. Scattering in a Central Field. Spinor Scattered Amplitude.

Let us go on to a consideration of the general theory of electron scattering in central fields. In order to find the differential cross section, we must find a solution to the Dirac equation whose asymptotic representation is a superposition of a plane wave and an outgoing spherical wave,

$$\psi \approx u e^{i p r} + G(n) \frac{e^{i p r}}{r}.$$

Here u is a unit bispinor which determines the polarization of the incident wave, the z axis is along the direction of the incident beam, and G is the amplitude of the scattered wave, which is also a bispinor and depends on the direction of scattering $\underline{n}$. The differential cross section is related to the amplitude G , as in nonrelativistic quantum mechanics, by the expression

$$d\sigma = |G|^2 d\Omega.$$

Asymptotically the Dirac equation for an electron in a field (which is, of course, assumed to vanish at infinity) is the same as the free-electron equation. Therefore, the relation between the two spinors which are the components of the wave function

$$\psi = \begin{pmatrix} \chi \\ \chi \end{pmatrix},$$

is the same as the previous one [see (8.12)], namely

$$\chi = \frac{p}{\epsilon + m} (\sigma n) \varphi,$$

where $\underline{n}$ is the direction of propagation of the wave at a given point. Introducing, in correspondence with (9.6), the notation

$$u = N \begin{pmatrix} v \\ w \end{pmatrix},$$

where $\underline{v}$ is a unit spinor, and similarly setting

$$G = N \begin{pmatrix} F \\ F' \end{pmatrix},$$

we can write

$$|G|^2 = \frac{|G|^2}{|u|^2} = \frac{F^2 \left[1 + \left(\frac{p}{\epsilon + m} \right)^2 (\sigma n)^2 \right]}{v^2 \left[1 + \left(\frac{p}{\epsilon + m} \right)^2 v^2 \right]} = \frac{|F|^2}{|v|^2} = |F|^2.$$

Thus, for the scattering problem it is sufficient to consider only part of the wave function, namely the spinor φ . If we write its asymptotic behavior in the form

$$\varphi = v e^{i p r} + F \frac{e^{i p r}}{r}, \quad (13.5)$$

then

$$d\sigma = |F|^2 d\Omega, \quad (13.6)$$

3. Scattering Cross Section in terms of the Phase Shifts.

To find an asymptotic solution of the form (13.5) for Equation (11.7), let us make use of the expansion (10.24) of a plane wave in terms of spherical waves:

$$v e^{i p r} = \sum_{j, l, m} (v \Omega_{j, l, m} \left(\frac{p}{r} \right)) g_{j, l, m}(p r) \Omega_{j, l, m} \left(\frac{r}{r} \right), \quad (13.7)$$

where $g_{j, l, m}$ is given by (4.29).

We note that if $\underline{p}$ is directed along the z axis, expression (13.7) is somewhat simplified; in this case

$$\Omega_{j, l, m} \left(\frac{p}{r} \right) = \Omega_{j, l, m}(0).$$

(The argument 0 is used to indicate that $\vartheta = 0$.)

According to the definition of the spherical spinor [see (10.6), (10.7)]

$$v \Omega_{j, l, m}(0) = \sum_{\lambda = -1/2}^{1/2} v(\lambda) \pi_{l, m}^{\lambda} Y_{l, m-1}(0),$$

and since the spherical function Y_{lm} is different from zero on the z axis only if $m = 0$,

$$Y_{lm}(0) = \sqrt{\frac{2l+1}{4\pi}} \delta_{m0},$$

we obtain

$$vQ_{jlm}^*(0) = \eta_{lm}^j v(M) \sqrt{\frac{2l+1}{4\pi}} \quad (13.8)$$

so that in (13.7) $\underline{M}$ can take on only the values $\pm \frac{1}{2}$. We shall emphasize this fact by replacing the summation index $\underline{M}$ by λ , writing

$$ve^{i\varphi} = \sum_{j\lambda} (vQ_{j\lambda}^*(0)) g_j(pr) Q_{j\lambda}\left(\frac{r}{r_0}\right). \quad (13.9)$$

Inserting the asymptotic expression for g_j into (13.9), we obtain

$$ve^{i\varphi} \approx 4\pi \sum_{j\lambda} i^l (vQ_{j\lambda}^*(0)) \frac{\cos\left(pr - (l+1)\frac{\pi}{2}\right)}{pr} Q_{j\lambda}. \quad (13.10)$$

According to (11.6) and (11.15) the general solution of the Dirac equation for an electron in an external field has the following asymptotic form:

$$\varphi = \sum_{j\lambda} C_{j\lambda}^M Q_{j\lambda} \frac{\cos\left(pr + \delta_{j\lambda} - (l+1)\frac{\pi}{2}\right)}{r}, \quad (13.11)$$

where the $C_{j\lambda}^M$ are arbitrary constants, and $\delta_{j\lambda} = \delta_{j\lambda} - (l+1)\frac{\pi}{2}$. Comparing this last expression with (13.10), we can easily find $C_{j\lambda}^M$ such that (13.11) is the same as (13.5). We achieve this by the choice

$$C_{j\lambda}^M = 4\pi i^l (vQ_{j\lambda}^*(0)) \frac{1}{p} e^{i\delta_{j\lambda}}.$$

Then the scattered wave amplitude is

$$F = \frac{4\pi}{p} \sum_{j\lambda} (vQ_{j\lambda}^*(0)) \frac{e^{2i\delta_{j\lambda}} - 1}{2} Q_{j\lambda}(\theta). \quad (13.12)$$

We obtain an expression for the total cross section by inserting (13.12) into (13.6) and making use of the orthonormal properties of the spherical spinors;

$$\sigma = \frac{(4\pi)^2}{p^2} \sum_{j\lambda} \sin^2 \delta_{j\lambda} \sum_{\lambda} |vQ_{j\lambda}^*(0)|^2.$$

Inserting (13.8) and noting that according to (10.7)

$$(\eta_{lm}^j)^2 = \frac{|x|}{2l+1},$$

is independent of λ , where κ is the quantum number defined in (11.8), and

$$\sum_{\lambda} |v(\lambda)|^2 = 1,$$

we finally obtain

$$\sigma = \frac{4\pi}{p^2} \sum_{\kappa} |x| \sin^2 \delta_{\kappa} \quad (\delta_{\kappa} = \delta_{j\lambda}). \quad (13.13)$$

Equation (13.13) becomes the nonrelativistic expression for the scattering cross section

$$\sigma = \frac{4\pi}{p^2} \sum_l (2l+1) \sin^2 \delta_l,$$

if the phase shifts δ_{κ} are the same for both values of κ

$$\kappa = \begin{cases} -l-1 \\ l \end{cases}$$

for a given l .

4. Azimuthal Asymmetry.

A significant difference between the results obtained above and the corresponding nonrelativistic ones is the fact that the differential cross section depends in general not only on the angle θ between the incident and scattered beams, but also on the azimuth angle φ . Equation (13.6) can be written in the form

$$d\sigma = P(\theta) [1 + \Delta(\theta) \cos(\varphi - \varphi_0)] d\varphi, \quad (13.14)$$

where P and Δ are functions only of θ . We shall now prove Equation (13.14) and shall give general expressions for P and Δ .

For this purpose let us express the spherical spinors in (13.12) in terms of spherical functions according to (10.6) and (10.7). By using (13.8), we obtain the components of the scattered amplitude F^{μ} in the form

$$F^{\mu} = \frac{2\pi}{ip} \sum_{\lambda} \sqrt{\frac{2l+1}{4\pi}} (e^{i\lambda\mu} - 1) \sum_{\lambda} \sigma(\lambda) \eta_{\lambda\mu}^{i\lambda} Y_{l,\lambda-\mu}.$$

Since λ and μ take on the values $\pm 1/2$, the spherical functions $Y_{l,m}$ in this expression have values of m given by

$$m = 0, \pm 1.$$

Therefore, the azimuth angle φ enters into F^{μ} in the form

$$m = 0, \pm 1.$$

Let us write $Y_{l,m}$ in terms of the associated Legendre polynomials P_l^m :

$$Y_{l0} = \sqrt{\frac{2l+1}{4\pi}} P_l^0(\theta),$$

$$Y_{l,\pm 1} = \mp \sqrt{\frac{2l+1}{4\pi}} \frac{1}{\sqrt{l(l+1)}} P_l^1(\theta) e^{\pm i\varphi}.$$

Using the explicit expressions for the coefficients $\eta_{\lambda\mu}^{i\lambda}$ [see (10.7)], we obtain

$$F^{\mu} = \frac{1}{2ip} \sum_{\lambda} (e^{2i\lambda} - 1) \left\{ \sigma\left(\frac{1}{2}\right) |x| P_l^0 - \sigma\left(-\frac{1}{2}\right) P_l^1 e^{-i\varphi} \frac{x}{|x|} \right\},$$

$$F^{-\mu} = \frac{1}{2ip} \sum_{\lambda} (e^{2i\lambda} - 1) \left\{ \sigma\left(-\frac{1}{2}\right) |x| P_l^0 + \sigma\left(\frac{1}{2}\right) \frac{x}{|x|} P_l^1 e^{i\varphi} \right\}. \quad (13.15)$$

Let us now introduce two functions of the angle θ , namely

$$\left. \begin{aligned} a(\theta) &= \frac{1}{2ip} \sum_{\lambda} (e^{2i\lambda} - 1) |x| P_l^0(\theta), \\ b(\theta) &= \frac{1}{2ip} \sum_{\lambda} (e^{2i\lambda} - 1) \frac{x}{|x|} P_l^1(\theta). \end{aligned} \right\} \quad (13.16)$$

Then (13.15) can be written

$$\left. \begin{aligned} F^{\mu} &= \sigma\left(\frac{1}{2}\right) a(\theta) - \sigma\left(-\frac{1}{2}\right) b(\theta) e^{-i\varphi}, \\ F^{-\mu} &= \sigma\left(-\frac{1}{2}\right) a(\theta) + \sigma\left(\frac{1}{2}\right) b(\theta) e^{i\varphi}. \end{aligned} \right\} \quad (13.17)$$

Inserting (13.17) into (13.12) and comparing with (13.14), we see that

$$\left. \begin{aligned} P(\theta) &= |a|^2 + |b|^2, \\ \Delta(\theta) \cos(\varphi - \varphi_0) &= \frac{a^*b - a b^*}{|a|^2 + |b|^2} \left[-\sigma\left(\frac{1}{2}\right) \sigma^*\left(-\frac{1}{2}\right) e^{i\varphi} + \sigma^*\left(\frac{1}{2}\right) \sigma\left(-\frac{1}{2}\right) e^{-i\varphi} \right]. \end{aligned} \right\} \quad (13.18)$$

Let us find those cases in which the azimuthal asymmetry does not exist ($\Delta = 0$). In the first place, according to (13.18), $\Delta = 0$, if $a(\theta)$ and $b(\theta)$ are real or if one of them vanishes. In particular, if δ_{κ} does not depend on the sign of κ , then according to (13.16) $b = 0$. This corresponds, for instance, to the nonrelativistic limit, as we have already seen. We note that then according to (13.17) the polarization of the beam is conserved, that is

$$\frac{F^{\mu}}{F^{-\mu}} = \frac{\sigma\left(\frac{1}{2}\right)}{\sigma\left(-\frac{1}{2}\right)}.$$

In the second place, $\Delta = 0$ if the initial beam is polarized parallel (or antiparallel) to the direction of motion, i.e., if $\underline{v} \cdot (-\underline{1}/2) = 0$ or $\underline{v} \cdot (\underline{1}/2) = 0$. Thus, there will be no azimuthal asymmetry for scattering of an unpolarized beam, since the scattering cross section for an unpolarized beam can be written as the sum of scattering cross sections for two beams whose polarizations are in the positive and negative $\underline{z}$ directions.

5. Polarization in Scattering.

Azimuthal asymmetry occurs when an electron beam polarized at some angle to the direction of motion is scattered. Such a beam can be obtained by scattering an unpolarized beam. The function $\Delta(\theta)$ which characterizes the azimuthal asymmetry of scattering can be expressed in terms of the functions a and b defined by (13.16).

In order to determine Δ , let us replace the unpolarized beam (as was mentioned above) by two beams for one of which $\chi(-1/2) = 0$, and for the other of which $\chi(1/2) = 0$. Let $\underline{y} = (\frac{1}{2}, \frac{1}{2})$. Then according to (13.17) the scattered amplitude in some direction $\underline{u}$ given by the angles $\theta_1, \varphi_1 = 0$ can be written in the form

$$F = \begin{pmatrix} a(\theta_1) \\ b(\theta_1) \end{pmatrix}.$$

Considering just the electrons scattered in this direction, we can describe them by a plane wave

$$\varphi_1 = \varphi_1 e^{i\varphi_1 r},$$

where $\underline{y}_1$ is the normalized amplitude

$$\varphi_1 = \frac{1}{\sqrt{|a(\theta_1)|^2 + |b(\theta_1)|^2}} \begin{pmatrix} a(\theta_1) \\ b(\theta_1) \end{pmatrix}.$$

In order to consider secondary scattering of this wave with the aid of the formulas derived above, we must transform the coordinates so that the new $\underline{z}$ axis (which we shall call the $\underline{z}'$ axis) is directed along $\underline{u}$. To do this we perform a rotation about the $\underline{y}$ axis by an angle θ_1 . Then according to the transformation rules for spinors, $\underline{y}_1$ becomes¹⁾

$$\varphi'_1 = \frac{1}{\sqrt{|a(\theta_1)|^2 + |b(\theta_1)|^2}} \begin{pmatrix} a(\theta_1) \cos \frac{\theta_1}{2} + b(\theta_1) \sin \frac{\theta_1}{2} \\ b(\theta_1) \cos \frac{\theta_1}{2} - a(\theta_1) \sin \frac{\theta_1}{2} \end{pmatrix}. \quad (13.19)$$

We can now make use of (13.18), setting $\underline{y} = \underline{y}'$, $\underline{a} = \underline{a}(\theta_2)$, $\underline{b} = \underline{b}(\theta_2)$, $\varphi = \varphi_2$, where θ_2 and φ_2 are the scattering angles from the second scattering center.

¹⁾ This transformation is easy to obtain with the aid of the fundamental transformation equation for spinors (7.2). For rotation about the $\underline{y}$ axis by a finite angle θ_1 , we integrate (7.2), obtaining

$$\varphi' = e^{-\frac{i}{2} \theta_1 \sigma_y} \varphi.$$

Writing the exponential in the form of a series and making use of the fact that $\sigma_y^2 = 1$, we obtain

$$\varphi' = \left(\cos \frac{\theta_1}{2} + i \sin \frac{\theta_1}{2} \sigma_y \right) \varphi,$$

which leads to (13.19).

Using the same method for the other beam, for which

$$\varphi = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad F = \begin{pmatrix} -b(\theta_1) \\ a(\theta_1) \end{pmatrix},$$

and taking the arithmetic average of the cross sections, we obtain

$$\Delta = -2 \frac{[a(\theta_1) b^*(\theta_1) - a^*(\theta_1) b(\theta_1)] [a(\theta_2) b^*(\theta_2) - a^*(\theta_2) b(\theta_2)]}{[|a(\theta_1)|^2 + |b(\theta_1)|^2] [|a(\theta_2)|^2 + |b(\theta_2)|^2]}. \quad (13.20)$$

6. Scattering in a Coulomb Field.

The above general theory of scattering in a central field can be applied to the Coulomb field. Since, however, the phase shift in the asymptotic expression for the wave function (see Section 12) contains a term logarithmic in r , we must attempt to find a solution of the Dirac equation in a form somewhat different from (13.5) to describe this scattering process, namely in the form

$$\varphi = \varphi_0 e^{i p r - i \gamma \ln p(r-r_0)} + F \frac{e^{-i p r - i \gamma \ln 2 p r}}{r}.$$

Then Equations (13.12) and (13.18) remain valid, and according to (12.24)²⁾

$$\delta_\kappa = \eta - \arg \Gamma(\gamma + i\eta) + \frac{\pi}{2} (l + 1 - \gamma). \quad (13.21)$$

In this case it is impossible to sum the series (13.16) for $a(\theta)$ and $b(\theta)$ analytically. For small values of Z it makes sense to write these functions as power series in $Z e^2$. Then up to terms of order $(Z e^2)^3$, the expression so obtained for the differential scattering cross section is identical with (13.4), which was obtained with the aid of perturbation theory. Here the functions $\underline{a}$ and $\underline{b}$ are real, so that no azimuthal asymmetry is present.

The next term of the series gives an expression for the scattering cross section which we shall present here without proof³⁾

$$d\sigma = \left(\frac{Z e^2}{2 p v} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}} \left[1 - v^2 \sin^2 \frac{\theta}{2} + \pi Z e^2 v \sin \frac{\theta}{2} \left(1 - \sin \frac{\theta}{2} \right) \right] d\theta. \quad (13.22)$$

²⁾ We note that according to the definition (12.19) of η , when $Z = 0$ we have $e^{i\eta} = -\frac{\kappa}{|\kappa|}$, so that δ_κ vanishes.

³⁾ See Mott and Massey, Theory of Atomic Collisions (Foreign Lit. Press, 1951).

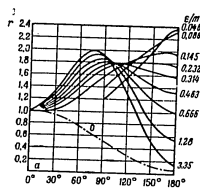


Fig. 2a. (Electrons)

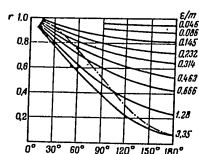


Fig. 2b. (Positrons)

As opposed to (13.4), Equation (13.22) gives different values for electrons and positrons. The expression for the positron scattering cross section can be obtained from the corresponding electron expression by replacing $\underline{z}$ by $-\underline{z}$. Therefore, the term $\underline{z}^2$ in (13.22) has different signs for electrons and positrons.

For the azimuthal asymmetry of double scattering this approximation gives the following expression when $\theta_1 = \theta_2 = \frac{\pi}{2}$:

$$\Delta = (Ze^2)^2 \frac{(1-v^2)v^2}{(2-v^2)^2}. \quad (13.23)$$

In general, for arbitrary Z the cross section must be calculated numerically. Figure 2 shows some results¹⁾ for mercury ($Z = 80$). The graphs give

$$r = \frac{\frac{d\sigma}{d\Omega}}{\left(\frac{Ze^2}{2pv \sin^2 \frac{\theta}{2}} \right)^2}$$

as a function of the scattering angle ϑ for various values of ϵ/m . Figure 3 gives the dependence of Δ on v (also for mercury) for $\vartheta_1 = \vartheta_2 = \frac{\pi}{2}$.

We shall also give some results⁹⁾ for high energies (in practice these results are applicable for energies greater than 4 Mev) when $\gamma \approx 1$. The differential cross section for this case can be written

$$\frac{d\sigma}{d\omega} = \left(\frac{Ze^2}{mv^2}\right)^2 A + \frac{1}{p^2} B,$$

²⁾ McKinley and Feshbach, Phys. Rev. 74, 1759 (1948). The dotted curve on Fig. 2 gives Equation (13.22) with $\epsilon/\bar{m} = 3.35$.
³⁾ H. Feshbach, Phys. Rev. 88, 295 (1952).

²⁰ H. Feshbach, Phys. Rev. 88, 295 (1952).

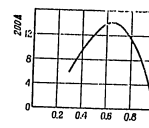


Fig. 3

where A and B are functions of the angle and of the nuclear charge and are independent of the energy. The values of these functions are given in the following table (negative values of Z refer to positron scattering).

For very large electron energies it is necessary to take account of the finite dimensions of the nucleus. The corresponding variation of the phase shifts can be calculated by the method indicated in Section 12. We shall not, however, present the corresponding results here ⁴⁾

*Values of A and B (the first number gives A, and the second B).

[illegible]

$Z \setminus \theta$		100°	120°	135°	150°		
-80		0.5309	0.0707	0.4263	0.0111	0.3514	0.0044
-70		0.4952	0.0438	0.2947	0.016	0.3055	0.0231
-60		0.4693	0.0268	0.2378	0.0038	0.2685	0.0004
-50		0.4446	0.0109	0.1810	0.0096	0.2291	0.0043
-40		0.4303	0.0023	0.3718	0.0099	0.3264	0.0017
-30		0.4310	0.0031	0.3803	0.0102	0.2973	0.0006
-20		0.4541	0.0019	0.3803	0.0076	0.3190	0.0034
0		0.5111	0.0704	0.4387	0.0281	0.3781	0.0032
20		0.5918	0.18	0.5099	0.0692	0.441	0.0329
40		0.7337	0.431	0.6987	0.168	0.6869	0.0935
60							0.6808
80							0.0137

¹⁾ See Rose, Phys. Rev. 73, 279 (1948); Feshbach, Phys. Rev. 84, 1206 (1951); Acheson, Phys. Rev. 82, 488 (1951).

§ 14. The Nonrelativistic Limit.

1. Transition to the Pauli Equation.

If the electron energy is not very different from the rest energy, that is if

$$|e - m| \ll m,$$

then as we have seen, one of the two spinors making up the electron wave function is much smaller than the other;

$$\chi \ll \varphi.$$

Making use of this fact we can construct an approximate equation which contains only the spinor φ . In this section we shall explicitly express the velocity of light by c (we have previously taken it as unity). Then the approximate equation can be obtained by an expansion in powers of $\frac{1}{c}$, since we assume that the velocity v of the electron is much smaller than the velocity of light

$$\frac{v}{c} \ll 1.$$

Let us start with the Dirac equation in an external field, which we shall write

$$i \frac{\partial \psi}{\partial t} = \left[c \alpha \left(p - \frac{e}{c} A \right) + \beta m c^2 + e A_0 \right] \psi. \quad (14.1)$$

Since in the nonrelativistic theory the energy is taken to mean the difference between the total energy and the rest energy, it is convenient to replace the wave function ψ by the function ψ' given by

$$\psi = \psi' e^{-i m c^2 t}. \quad (14.2)$$

Inserting (14.2) into (14.1), we obtain

$$i \frac{\partial \psi'}{\partial t} = \left\{ c \alpha \left(p - \frac{e}{c} A \right) + (\beta - 1) m c^2 + e A_0 \right\} \psi'. \quad (14.3)$$

Let us write ψ' , as usual, in the form

$$\psi' = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$$

[we have dropped the primes on φ and χ ; this can lead to no confusion, since we shall henceforth make use only of the function ψ' defined by (14.2)]. Equation (14.3) is equivalent to the following set of equations:

$$i \frac{\partial \varphi}{\partial t} = c \alpha \left(p - \frac{e}{c} A \right) \chi + e A_0 \varphi, \quad (14.4)$$

$$i \frac{\partial \chi}{\partial t} = c \alpha \left(p - \frac{e}{c} A \right) \varphi + e A_0 \chi - 2 m c^2 \chi. \quad (14.5)$$

In order to expand this in powers of $\frac{1}{c}$, let us assume (this shall be verified by what follows) that χ is of order of magnitude $\frac{1}{c}$ φ . Then we may drop all terms containing χ in (14.5), except the last, whose coefficient is proportional to c^2 . We thus obtain

$$\chi = \frac{1}{2 m c} \sigma \left(p - \frac{e}{c} A \right) \varphi. \quad (14.6)$$

Inserting (14.6) into (14.4) we arrive at

$$i \frac{\partial \varphi}{\partial t} = \left\{ \frac{1}{2 m} \left[\sigma \left(p - \frac{e}{c} A \right) \right]^2 + e A_0 \right\} \varphi.$$

Using the properties of the σ_i matrices given by (7.6), it is easily shown that

$$\left[\sigma \left(p - \frac{e}{c} A \right) \right]^2 = \left(p - \frac{e}{c} A \right)^2 - \frac{e}{c} \sigma H,$$

where $H = \text{curl } A$ is the magnetic field. With this relation we can transform the first order equation into the form

$$i \frac{\partial \varphi}{\partial t} = H \varphi,$$

where

$$H = \frac{1}{2m} \left(p - \frac{e}{c} A \right)^2 + eA_0 - \frac{e}{2mc} \sigma H. \quad (14.7)$$

The Schrodinger equation with the Hamiltonian given by (14.7) is called the Pauli equation. It differs from the nonrelativistic Schrodinger equation by the last term in the Hamiltonian, which has the form of the potential energy of a magnetic dipole in the external field. Thus, in the first approximation the electron behaves as a nonrelativistic particle which possesses, in addition to its charge, a magnetic moment

$$\mu = \frac{eh}{2mc} \sigma. \quad (14.8)$$

2. Second Approximation.

Let us go on to the second approximation, taking terms up to those of order $\frac{1}{c^2}$ in the expansion. We shall restrict ourselves to the case in which there is only an external electric field ($A = 0$).

Let us replace χ in the terms we have previously dropped from Equation (14.5) by its first-approximation expression (see (14.6)), and then let us solve (14.5) for χ . We then obtain

$$\chi = \frac{1}{2mc} \left(\sigma p + \frac{eA_0}{2mc^2} \sigma p - \frac{i}{2mc^2} \sigma p \frac{\partial}{\partial t} \right) \varphi.$$

Now inserting this second-order expression for χ into (14.4), we obtain an equation for φ , namely

$$i \left(1 + \frac{p^2}{4m^2c^2} \right) \frac{\partial \varphi}{\partial t} = \left\{ \frac{p^2}{2m} + eA_0 + \frac{e}{4m^2c^2} (\sigma p) A_0 (\sigma p) \right\} \varphi.$$

Multiplying this by $(1 - \frac{p^2}{4m^2c^2})$ and maintaining terms of order $\frac{1}{c^2}$ and lower, we have

$$\left. \begin{aligned} i \frac{\partial \varphi}{\partial t} &= H' \varphi, \\ H' &= \left(1 - \frac{p^2}{4m^2c^2} \right) \left(\frac{p^2}{2m} + eA_0 \right) + \frac{e}{4m^2c^2} (\sigma p) A_0 (\sigma p). \end{aligned} \right\} \quad (14.9)$$

This equation cannot, however, be treated directly as a Schrodinger equation. The function φ entering into (14.9) does not have exactly the meaning of an electron wave function, and the operator H' acting on φ cannot be considered a Hamiltonian. Indeed, in the first place the integral $\int \varphi^* \varphi dr$ is not constant in

time (according to the continuity equation, $\int (\varphi^* \varphi + \chi^* \chi) dr$ is conserved). Thus, the function φ cannot be normalized. Furthermore, the operator H' in (14.9) is not Hermitian (the non-Hermitian operator is $-\frac{e}{4m^2c^2} p^2 A_0$). Both of these failings are removed if we transform the wave function, introducing instead of φ , the wave function Φ given by

$$\Phi = O \varphi \quad (14.10)$$

where O is chosen so that Φ can be normalized.

With (14.10), Equation (14.9) becomes

$$i \frac{\partial \Phi}{\partial t} = H \Phi,$$

where

$$H = O H' O^{-1}. \quad (14.11)$$

The operator O can be found from the requirement that up to terms of order $\frac{1}{c^2}$

$$\int (\varphi^* \varphi + \chi^* \chi) dr = \int \Phi^* \Phi dr.$$

It is sufficient to use the first approximation for χ as given by (14.6) in this equation, so that to the order demanded we have

$$\int \varphi^* \left(1 + \frac{p^2}{4m^2c^2} \right) \varphi dr = \int \Phi^* \Phi dr. \quad (14.12)$$

It follows from (14.12) that the operator O can be chosen as

$$O = \left(1 + \frac{p^2}{8m^2c^2} \right), \quad (14.13)$$

whence

$$O^{-1} = \left(1 - \frac{p^2}{8m^2c^2} \right).$$

Since the operator O differs from unity by a term of order $\frac{1}{c^2}$, the transformation (14.10) does not alter the first order equation; in which ψ is a normalized wave function up to terms of order $\frac{1}{c^2}$ and the operator of (14.7) is a Hamiltonian.

Let us now find the Hamiltonian operator in the second approximation. Inserting expressions (14.13) and (14.9) for O and H' into (14.11), we obtain, up to terms of order $\frac{1}{c^2}$,

$$H = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + eA_0 + \frac{1}{8m^2c^2} (p^2 A_0 - A_0 p^2) - \frac{e}{4m^2c^2} p^2 A_0 + \frac{e}{4m^2c^2} (\sigma p) A_0 (\sigma p). \quad (14.14)$$

The last two terms of (14.14) contain non-Hermitian operators. It is, however, easy to show that they cancel. Indeed, the second term can be written

$$(\sigma p) A_0 (\sigma p) = A_0 p^2 - \sigma (A_0 p - p A_0) \sigma p = A_0 p^2 - i(\sigma \nabla A_0) (\sigma p) = A_0 p^2 + iE p - [E p] \sigma,$$

where $\mathbf{E} = -\nabla A_0$ is the electric field. Inserting this into (14.14) and noting that

$$p^2 A_0 - A_0 p^2 = -\Delta A_0 + 2iE p,$$

we obtain the operator H in the form

$$H = \frac{p^2}{2m} + eA_0 - \frac{p^4}{8m^3c^2} - \frac{e}{4m^2c^2} \sigma [E p] + \frac{e}{8m^2c^2} \Delta A_0, \quad (14.15)$$

which can be easily seen to be Hermitian.

The last three terms of (14.15) are corrections of order $\frac{1}{c^2}$. The first of these gives the dependence of the mass on the velocity, and the second can be interpreted as the interaction between a moving magnetic dipole and an electric field. Since for an electrostatic field

$$\Delta A_0 = -4\pi\rho,$$

where ρ is the charge density which produces the external field, the last term differs from zero only at points where charges are located. For the Coulomb field of the nucleus,

$$\Delta A_0 = 4\pi Ze\delta(r).$$

With the aid of perturbation theory and (14.15) we can find the hydrogen fine structure [the perturbation in this case consists of the three last terms in (14.15)]. This gives Equation (12.15), which was obtained by a power series expansion of the exact solution (12.14) for the Dirac equation.

The second approximation for the case in which there exists an external magnetic field can be obtained by the same method as was used to obtain (14.15). The result differs from (14.15) in that p is replaced by $p - \frac{e}{c} \mathbf{A}$, and $-p^2$ by $[\sigma(p - \frac{e}{c} \mathbf{A})]^2 = (p - \frac{e}{c} \mathbf{A})^2 - \frac{e}{c} \sigma \mathbf{H}$.

3. On the Applicability of the Dirac Equation to Nucleons.

In principle the Dirac equation is applicable to any particle with spin $\frac{1}{2}$. In practice, however, it can be actually used only for the study of electron motion. This is explained by the fact that the fundamental interaction for the electron is the electromagnetic one. For the proton, on the other hand, meson field interactions and the nuclear forces related to them are of great significance. In view of the absence of a satisfactory theory of interaction with these meson fields, the applicability of the Dirac equation to the proton is quite limited.

Nor, even, can the behavior of the proton in an electromagnetic field be investigated in a satisfactory way. Due to the strong interaction with the meson field, which is related to particles having electric charge, the motion of the proton in an electromagnetic field is not the same as that of a particle having only a charge. Therefore, Equation (11.1) is not applicable to the proton; the exact equation would have to take account of the additional interaction with the electromagnetic field by means of the meson fields.

In particular, as has been shown by experiment, Equation (14.7) of the nonrelativistic limit cannot be used for the proton. When $v \ll c$ the proton has a magnetic moment not given by $\mu = \frac{e}{2M}$ (where M is the proton mass), as would follow from (14.8), but one given by $\mu_p = 2.8\mu$. The neutron, with no charge, has a magnetic moment $\mu_n = -1.9\mu$. The additional "anomalous" magnetic moment of nucleons can be treated by inserting in the Dirac equation a further interaction of the form

$$-(\mu - \mu_0) \mathbf{H}.$$

The corresponding Dirac equation would then be

$$(\hat{t}_4 p_4 - ie\hat{t}_1 A_1 - \hat{t}_2' \hat{t}_3' F_{12} + m) \psi = 0,$$

where F_{ik} is the electromagnetic field tensor, and $\mu' = \mu_p - \mu$. However, the introduction of the anomalous magnetic moment of the nucleon has only limited meaning and is applicable only in the region of low particle energies and low electromagnetic field frequencies.

When in the future we make use of the Dirac equation to describe nucleons (Sections 30, 39) we shall bear in mind this restriction on its applicability. Then the Dirac equation is equivalent to (14.7) with μ replaced by μ_p .

In the previous sections we have set forth a theory of photons and electrons on the basis of particle quantum mechanics. We have seen, however, that the fundamental concepts of quantum mechanics must be altered significantly when considering particles moving at high velocities (in the relativistic region). Thus, the potentials of the field corresponding to the photon cannot be considered the usual wave functions of quantum mechanics. As has already been noted in Section 1, the wave function of the photon can be defined only in momentum space. The meaning of the electron wave function also changes in the relativistic region. As we have seen (Section 8) different solutions of the electron wave equation (the Dirac equation) must be assigned to different particles (the electron or positron). Thus, the interpretation of the wave function as the amplitude of the probability density

for localizing a particle does not, strictly speaking, exist in relativistic quantum mechanics. Actually, the quantum mechanics of an individual particle cannot be consistently formulated. This is closely related to the fact that in reality the number of particles is not conserved in the process of motion (pair creation, radiation, etc.) due to the various interactions.

On the other hand, the value of the concept of a field corresponding to a particle is hardly reduced; on the contrary, it increases. We shall see later that the consistent quantum theory of fields describes the phenomena of production, annihilation, and transformation of particles, and that there exist states of the field in which no particles are present (vacuum states).

CHAPTER III

QUANTIZATION OF THE ELECTROMAGNETIC AND ELECTRON-POSITRON FIELDS

§ 15. Quantization of the Electromagnetic Field.

1. Four-Dimensional Form of the Field Equations. Plane Waves.

Let us return to an examination of the properties of the free electromagnetic field.

In Chapter I we saw how to construct the wave function of the photon and to consider its quantum states on the basis of the classical Maxwell' equations. By developing the method of momentum-space wave functions, we can construct the wave function of an arbitrary number of photons, as well as generalize this to a state with an undetermined number of photons (see, in this respect, paragraphs 4 and 5 of this Section).

For a general formulation of the theory of the free electromagnetic field and its interaction with electrons, as well as for the investigation of concrete phenomena due to this interaction, it is more convenient to make use of a different method,¹⁾ which we shall now describe. It consists of considering the electromagnetic field as a generalized dynamic system, and the variables which determine this field as the generalized coordinates of the system. The quantum nature of this system is accounted for by replacing the classical quantities by the operators corresponding to them. These operators operate on the wave function of the system, which depends on the "occupation numbers," which are the numbers of particles in the various quantum states.

The transition from the configuration-space (or momentum-space) wave function to the wave function in the space of occupation numbers is a canonical transformation (the method employing the wave function in the space of occupation numbers, and the procedure by which one transforms to this wave function, is usually called second quantization). Therefore, both approaches to the quantum theory of electromagnetic processes from the "corpuscular" (the photons and their wave functions) and "field" (the field and its quantum mechanical investigation) points of view are equivalent.

Thus, we shall no longer consider the potentials $\underline{A}$ and $\underline{A}_0$ and the electromagnetic field vectors $\underline{E}$ and $\underline{H}$ as classical quantities (c-numbers), but shall treat them as operators (q-numbers). Similarly, as the nonrelativistic quantum mechanics of a point mass preserves Newton's equations of motion in the relation between the position and momentum operators, Maxwell's equations for the field operators are preserved in quantum electrodynamics. Just as the equations of motion take on actual meaning in quantum mechanics only after the commutation relations between the displacements and momenta have been established, it shall first be necessary, in order to develop quantum electrodynamics, to establish the commutation relations between the dynamical variables which characterize the electromagnetic field.

Setting up the commutation rules satisfied by the field operators is called quantizing the field. These relations obviously should be invariant with respect to Lorentz transformation. Therefore, in setting them up, one should start from Maxwell's equations in their four-dimensional covariant form.

The electric and magnetic fields $\underline{E}$ and $\underline{H}$ which satisfy Maxwell's equations (1.3), can be combined to form the antisymmetric electromagnetic field tensor $F_{\mu\nu}$, whose components

¹⁾ This method was first developed by Dirac [Proc. Roy. Soc. A 114, 243, 710 (1927)] in application to the theory of radiation.

for localizing a particle does not, strictly speaking, exist in relativistic quantum mechanics. Actually, the quantum mechanics of an individual particle cannot be consistently formulated. This is closely related to the fact that in reality the number of particles is not conserved in the process of motion (pair creation, radiation, etc.) due to the various interactions.

On the other hand, the value of the concept of a field corresponding to a particle is hardly reduced; on the contrary, it increases. We shall see later that the consistent quantum theory of fields describes the phenomena of production, annihilation, and transformation of particles, and that there exist states of the field in which no particles are present (vacuum states).

CHAPTER III

QUANTIZATION OF THE ELECTROMAGNETIC AND ELECTRON-POSITRON FIELDS

§ 15. Quantization of the Electromagnetic Field.

1. Four-Dimensional Form of the Field Equations. Plane Waves.

Let us return to an examination of the properties of the free electromagnetic field.

In Chapter I we saw how to construct the wave function of the photon and to consider its quantum states on the basis of the classical Maxwell equations. By developing the method of momentum-space wave functions, we can construct the wave function of an arbitrary number of photons, as well as generalize this to a state with an undetermined number of photons (see, in this respect, paragraphs 4 and 5 of this Section).

For a general formulation of the theory of the free electromagnetic field and its interaction with electrons, as well as for the investigation of concrete phenomena due to this interaction, it is more convenient to make use of a different method,¹⁾ which we shall now describe. It consists of considering the electromagnetic field as a generalized dynamic system, and the variables which determine this field as the generalized coordinates of the system. The quantum nature of this system is accounted for by replacing the classical quantities by the operators corresponding to them. These operators operate on the wave function of the system, which depends on the "occupation numbers," which are the numbers of particles in the various quantum states.

The transition from the configuration-space (or momentum-space) wave function to the wave function in the space of occupation numbers is a canonical transformation (the method employing the wave function in the space of occupation numbers, and the procedure by which one transforms to this wave function, is usually called second quantization). Therefore, both approaches to the quantum theory of electromagnetic processes from the "corpuscular" (the photons and their wave functions) and "field" (the field and its quantum mechanical investigation) points of view are equivalent.

Thus, we shall no longer consider the potentials A and A_0 and the electromagnetic field vectors $\mathbf{E}$ and $\mathbf{H}$ as classical quantities (c-numbers), but shall treat them as operators (q-numbers). Similarly, as the nonrelativistic quantum mechanics of a point mass preserves Newton's equations of motion in the relation between the position and momentum operators, Maxwell's equations for the field operators are preserved in quantum electrodynamics. Just as the equations of motion take on actual meaning in quantum mechanics only after the commutation relations between the displacements and momenta have been established, it shall first be necessary, in order to develop quantum electrodynamics, to establish the commutation relations between the dynamical variables which characterize the electromagnetic field.

Setting up the commutation rules satisfied by the field operators is called quantizing the field. These relations obviously should be invariant with respect to Lorentz transformation. Therefore, in setting them up, one should start from Maxwell's equations in their four-dimensional covariant form.

The electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$ which satisfy Maxwell's equations (1.3), can be combined to form the antisymmetric electromagnetic field tensor $F_{\mu\nu}$, whose components

¹⁾ This method was first developed by Dirac [Proc. Roy. Soc. A 114, 243, 710 (1927)] in application to the theory of radiation.

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & H_x & -H_y & -iE_z \\ -H_x & 0 & H_z & -iE_y \\ H_y & -H_z & 0 & -iE_x \\ iE_z & iE_y & iE_x & 0 \end{pmatrix}, \quad (15.1)$$

are related to the four-dimensional potential A_μ , with components (A, iA_3) by the expression

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}. \quad (15.2)$$

The first pair of Equations (1.3) is equivalent to (15.2). Indeed, differentiating (15.2) by x_ρ , we arrive at the four equations

$$\frac{\partial F_{\mu\alpha}}{\partial x_\rho} + \frac{\partial F_{\rho\alpha}}{\partial x_\mu} + \frac{\partial F_{\mu\rho}}{\partial x_\alpha} = 0, \quad (15.3)$$

which can be combined with (15.1) to be written in the form of (1.3).

The second pair of Equations (1.3) says that the four-dimensional divergence of the tensor $F_{\mu\nu}$ vanishes in the vacuum³⁾

$$\frac{\partial F_{\mu\nu}}{\partial x_\nu} = 0, \quad \mu = 1, 2, 3, 4. \quad (15.4)$$

This can be seen by inserting the components of $F_{\mu\nu}$ from (15.1) into (15.4). We note that Equation (15.4) is the same as (52.11) if the rest mass of the particle involved is zero. Therefore, the photon rest mass is zero.

Related to the fact that the photon rest mass is zero is the fact that it is only the field tensor $F_{\mu\nu}$ which has direct physical meaning, and not the four-dimensional potential A_μ which does not enter directly into Maxwell's equations (1.3); this is different than for the case $m \neq 0$ [see Equations (52.11)].

The potential A_μ is not uniquely defined by $F_{\mu\nu}$, since the gradient of an arbitrary scalar η can be added to A_μ without altering the field. In other words Maxwell's equations are invariant under the transformation

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{\partial \eta}{\partial x^\mu}, \quad (15.5)$$

where η is an arbitrary scalar. This transformation is called a gauge transformation.

All physical relations containing the electromagnetic field should, obviously, be invariant under gauge transformations.

Since the potential A_μ is not uniquely determined, we may require that it satisfy the subsidiary condition

$$\frac{\partial A_\mu}{\partial x_\mu} = 0. \quad (15.6)$$

Using Maxwell's equations (15.4), the definition of the field tensor (15.2), and Condition (15.6), it is easy to show that the components of the potential satisfy D'Alembert's equation

$$\square A_\mu = \frac{\partial^2}{\partial x_\mu^2} A_\mu = 0. \quad (15.7)$$

Together with the subsidiary condition (15.6), this is equivalent to Maxwell's equations.

The condition (15.6) does not uniquely determine the potential A_μ but merely restricts the group of gauge transformations to those functions η which satisfy the equation

$$\square \eta = 0. \quad (15.8)$$

It is important to note that in order for (15.6) to be satisfied, it is sufficient to demand that at some initial time $t = 0$, the quantity $\chi = \frac{\partial A_\mu}{\partial x_\mu}$ and its time derivative $\frac{\partial \chi}{\partial t}$ vanish. Then χ will remain zero for all time. Indeed, expanding χ in a power series in t , we have

$$\chi = \chi_0 + t \left(\frac{\partial \chi}{\partial t} \right)_0 + \frac{t^2}{2} \left(\frac{\partial^2 \chi}{\partial t^2} \right)_0 + \dots$$

(the index zero indicates that the derivatives are taken at the point $t = 0$). The first two terms in this series vanish by assumption. It follows from (15.7) that

$$\square \chi = 0,$$

and, therefore, that

$$\left(\frac{\partial^2 \chi}{\partial t^2} \right)_0 = \Delta \chi_0 = 0,$$

³⁾ Here, as in the future, the summation convention from $\nu = 1$ to $\nu = 4$ is used for indices that occur twice.

which means that the third term of the series vanishes. Similarly, it can be shown that all the rest of the terms in the expansion for χ also vanish, which means that $\chi = 0$. We thus see that the subsidiary condition (15.6) is the same as the initial conditions

$$\left(\frac{\partial A_\mu}{\partial x_\mu}\right)_0 = 0, \quad \left(\frac{\partial}{\partial t} \frac{\partial A_\mu}{\partial x_\mu}\right)_0 = 0, \quad (15.6')$$

The electromagnetic field is a real quantity. According to the general theory of wave fields (see Section 49) it follows from this that the particles related to the electromagnetic field, that is photons, are electrically neutral.

Maxwell's equation can be obtained from the variational principle if the Lagrangian density is chosen as

$$L = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} \quad (15.9)$$

It can be shown (see Section 49) that the energy-momentum tensor of the field is then given by the expression

$$T_{\mu\nu} = F_{\mu\alpha} F_{\nu\alpha} - \frac{1}{4} \delta_{\mu\nu} F_{\alpha\beta}^2 \quad (15.10)$$

The energy density of the electromagnetic field is¹⁾

$$w = -T_{44} = \frac{1}{2} (E^2 + H^2). \quad (15.11)$$

We note that the wave equations (15.7) for the potential can also be obtained from a variational principle, but then it is necessary to choose the Lagrangian not as (15.9), but as

$$L = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \frac{1}{2} \left(\frac{\partial A_\mu}{\partial x_\mu} \right)^2,$$

or

$$L = -\frac{1}{2} \frac{\partial A_\mu}{\partial x_\mu} \frac{\partial A_\mu}{\partial x_\mu} \quad (15.9')$$

¹⁾ Here the electric and magnetic fields are expressed in Heaviside units. In order to obtain the same expression in Gaussian units, $\mathbf{E}$ and $\mathbf{H}$ should be divided by $\sqrt{4\pi}$.

(these expressions differ by a divergence). In order now to obtain Maxwell's equations, we must add the subsidiary condition (15.6) to the wave equations.

If the Lagrangian is picked according to (15.9'), we obtain the following expression for the energy-momentum tensor:

$$T_{\mu\nu} = -\frac{\partial L}{\partial A_\mu} \frac{\partial A_\nu}{\partial x_\nu} + L \delta_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} \frac{\partial A_\mu}{\partial x_\nu} - \frac{1}{2} \delta_{\mu\nu} \frac{\partial A_\alpha}{\partial x_\alpha} \frac{\partial A_\alpha}{\partial x_\alpha}. \quad (15.10')$$

Henceforth, we shall use (15.7) as the wave equation with the subsidiary condition (15.6).

Let us expand the potential $A_\mu(\mathbf{x})$ at the point $\mathbf{x}(\mathbf{x}, t)$ in plane waves

$$A_\mu(x) = \frac{i}{\sqrt{V}} \sum_{\mathbf{k}, \lambda} (b_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + b_{\mathbf{k}\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{x}}) \varepsilon_{\mu\lambda}, \quad (15.12)$$

where $\mathbf{k}\cdot\mathbf{x} = \mathbf{k}\cdot\mathbf{x} - \omega t$ (for $\omega > 0$) is the scalar product of the four-dimensional position vector x and the four-dimensional wave vector $\mathbf{k}$ with components $k_\mu(k_\mu, i\omega)$, whose square magnitude is zero according to (15.7); $b_{\mathbf{k}\lambda}$ and $b_{\mathbf{k}\lambda}^*$ ($\lambda = 1, 2, 3, 4$) are the amplitudes, and the $\varepsilon_{\mu\lambda}$ are four-dimensional unit polarization vectors which determine the various vibrations for each wave vector $\mathbf{k}$; V is the normalizing volume. The summation is taken over all values of $\mathbf{k}$ and the four values of λ for each $\mathbf{k}$.

The unit polarization vectors are conveniently introduced in the following way. Let us consider some definite wave vector $\mathbf{k}$ and associate with it a coordinate system one of whose space axes (the 3 axis is directed along $\mathbf{k}$, the other two (the 1 and 2 axes) being perpendicular to $\mathbf{k}$ and to each other; the fourth (time) axis is the same as that which defines $k_4 = i\omega$ and $A_4 = i\phi$). We shall indicate these axes by the index $\lambda = \lambda(\mathbf{k})$, which takes on the values $\lambda = 1, 2, 3, 4$. The basis vectors corresponding to these axes shall be denoted by $\varepsilon_{\mu\lambda}$, and we shall consider $\varepsilon_{\mu\lambda}$ real. The vectors $\varepsilon_{\mu\lambda}$ satisfy the following orthonormality conditions:

$$\varepsilon_{\mu\lambda} \varepsilon_{\mu\lambda'} = \delta_{\lambda\lambda'}, \quad \lambda, \lambda' = 1, 2, 3, 4. \quad (15.13)$$

The scalar products of the $\varepsilon_{\mu\lambda}$ with $\mathbf{k}$ are given, obviously, by

$$\varepsilon_{\mu\lambda} k_\mu = \begin{cases} 0, & \lambda = 1, 2, \\ i\omega, & \lambda = 3, 4. \end{cases} \quad (15.14)$$

We note also the self-evident relation

²⁾ The symbol $b_{\mathbf{k}\lambda}^*$ denotes the complex conjugate of $-b_{\mathbf{k}\lambda}$, not of $+b_{\mathbf{k}\lambda}$; then $A_4 = iA_\phi$ will be pure imaginary [see Equation (15.14)].

$$e_{\lambda\mu}e_{\lambda\nu} = \delta_{\mu\nu}. \quad (15.15)$$

The potential, as we have seen before, must be subjected to the subsidiary condition (15.6). Inserting the expansion (15.12) into (15.6), and making use of (15.14), we obtain

$$\begin{aligned} \frac{\partial A_\mu}{\partial x_\mu} &= \frac{i}{\sqrt{V}} \sum_{k,\lambda} (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx}) e_{\mu\lambda} k_\mu = \\ &= \frac{i}{\sqrt{V}} \sum_k \omega (b_{k3} + ib_{k4}) e^{ikx} - (b_{k3}^* + ib_{k4}^*) e^{-ikx} = 0, \end{aligned}$$

whence we arrive at the following conditions for all k :

$$b_{k3} + ib_{k4} = 0, \quad b_{k3}^* + ib_{k4}^* = 0. \quad (15.16)$$

Let us now determine the field tensor. Making use of (15.2) and (15.12), we obtain

$$F_{\mu\nu} = \frac{i}{\sqrt{V}} \sum_{k,\lambda} (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx}) (e_{\nu\lambda} k_\mu - e_{\mu\lambda} k_\nu). \quad (15.17)$$

The electric and magnetic fields are given by

$$\left. \begin{aligned} E &= \frac{i}{\sqrt{V}} \sum_{k,\lambda=1,2} \omega (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx}) e_{\lambda 1} + \\ &+ \frac{i}{\sqrt{V}} \sum_k k \{ (b_{k3} - b_{k0}) e^{ikx} - (b_{k3}^* - b_{k0}^*) e^{-ikx} \} = \\ &= \frac{i}{\sqrt{V}} \sum_{k,\lambda=1,2} \omega (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx}) e_{\lambda 1} \\ H &= \frac{i}{\sqrt{V}} \sum_{k,\lambda=1,2} [k e_{\lambda 1} (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx})] \end{aligned} \right\} \quad (15.17')$$

where $\underline{e}_\lambda$ is the space part of $e_{\lambda\mu}$, and $b_{k0} = \frac{1}{i} b_{k4}$.

We see that due to the subsidiary condition (15.6)¹⁾ the electric and magnetic fields do not include oscillations with $\lambda = 3, 4$ (the so-called longitudinal and scalar vibrations) and are, therefore, transverse.

Let us determine the energy and momentum of the field. The energy-momentum tensor, according to (15.10*) and (15.17) can be written

$$T_{\mu\nu} = -\frac{1}{V} \sum_{k,\lambda; k',\lambda'} (b_{k\lambda} e^{ikx} - b_{k\lambda}^* e^{-ikx}) (b_{k'\lambda'} e^{ik'x} - b_{k'\lambda'}^* e^{-ik'x}) \times \\ \times \left(k_\mu k'_\nu - \frac{1}{2} \delta_{\mu\nu} k_\mu k'_\mu \right) e_{\lambda\lambda'}. \quad (15.18)$$

The energy and momentum of the field are defined by the equations

$$\mathcal{E} = - \int T_{44} d^3x, \quad P_j = \frac{1}{T} \int T_{4j} d^3x, \quad j = 1, 2, 3. \quad (15.19)$$

Using (15.18) it is easily shown that

$$\mathcal{E} = \sum_{k,\lambda=1,2} \omega^2 (b_{k\lambda} b_{k\lambda}^* + b_{k\lambda}^* b_{k\lambda}), \quad (15.20)$$

$$P = \sum_{k,\lambda=1,2} \omega k (b_{k\lambda} b_{k\lambda}^* + b_{k\lambda}^* b_{k\lambda}). \quad (15.21)$$

We see that the energy and momentum of the field, as could have been expected, do not depend on the longitudinal and scalar vibrations corresponding to $\lambda = 3, 4$, being determined entirely by the transverse vibrations (with $\lambda = 1, 2$).

2. Quantum Conditions.

Let us move on to establish the quantization conditions of the electromagnetic field. We shall start with the plane wave expansion (15.12) of the potential $A_\mu(\underline{x})$. Let us introduce instead of $b_{k\lambda}$ and $b_{k\lambda}^*$, the operators $c_{k\lambda}$ and $c_{k\lambda}^+$ defined by

$$b_{k\lambda} = \sqrt{\frac{\hbar}{2\omega}} c_{k\lambda}, \quad b_{k\lambda}^* = \sqrt{\frac{\hbar}{2\omega}} c_{k\lambda}^+ \quad (15.22)$$

and let us assume that $c_\lambda = c_{k\lambda}$, and $c_\lambda^+ = c_{k\lambda}^+$ satisfy the following commutation relations²⁾

¹⁾ The fundamental equations of the field are taken as the wave equation (15.7) with the subsidiary condition (15.6).
²⁾ The same relations for all values of λ insure, as we shall see later, the relativistic invariance of $[A_\lambda(\underline{x}), A_\mu(\underline{x}')]_-$.

$$[c_\lambda, c_\mu] = 0, \quad [c_\lambda^\dagger, c_\mu^\dagger] = 0, \quad [c_\lambda, c_\mu^\dagger] = \delta_{\lambda\mu}, \quad \lambda, \mu = 1, 2, 3, 4, \quad (15.23)$$

where $[AB] = AB - BA$.

We shall assume, to correspond to the fact that A is real, that the operators c_λ and $c_\lambda^\dagger$ for $\lambda = 1, 2, 3$ are Hermitian conjugates. Then it follows from the quantum conditions (15.23) that the eigenvalues of the operator $c_\lambda^\dagger c_\lambda$ ($\lambda = 1, 2, 3$), which we shall denote by N_λ , are the nonnegative integers $N_\lambda = 0, 1, 2, \dots$ ($\lambda = 1, 2, 3$).

The operator which is the Hermitian conjugate of $c_4 = ic_3$ is $-c_4^\dagger = -ic_3^\dagger$ (c_3 and $c_3^\dagger$ are Hermitian conjugates, since $A_4 = iA_3$ is pure imaginary). Thus, the quantum condition (15.23) for $\lambda = 4$ can be written

$$c_3^\dagger c_3 - c_4 c_4^\dagger = 1, \quad (15.23')$$

whence it follows that the eigenvalues of $c_3 c_3^\dagger$, which we shall call N_4 , are $N_4 = 0, 1, 2, \dots$

In the representation in which the $c_\lambda^\dagger c_\lambda$ are diagonal, the only matrix elements of c_λ and $c_\lambda^\dagger$ which do not vanish are¹⁾

$$\left. \begin{aligned} (c_\lambda)_{N_\lambda, N_\lambda+1} &= \sqrt{N_\lambda+1}, \quad (c_\lambda^\dagger)_{N_\lambda, N_\lambda-1} = \sqrt{N_\lambda} \quad (\lambda = 1, 2, 3), \\ (c_3^\dagger)_{N_4, N_4+1} &= \sqrt{N_4+1}, \quad (c_4)_{N_4, N_4-1} = \sqrt{N_4}. \end{aligned} \right\} \quad (15.24)$$

We shall now show that the quantum conditions give the correct corpuscular description. For this purpose let us determine the operators for the total energy and total momentum of the field in a volume V , and let us find their eigenvalues.

The energy and momentum operators are defined by (15.19) and (15.10'), in which the $F_{\mu\nu}$ are the operators

$$F_{\mu\nu} = -\frac{\partial}{\partial x_\mu} A_\nu + \frac{\partial}{\partial x_\nu} A_\mu,$$

where the A_μ are given by (15.12) and the quantum conditions (15.23).

¹⁾ This is easy to see by calculating $[c_\lambda^\dagger c_\lambda, c_\lambda^\dagger]$ and making use of (15.23) and the diagonality of $c_\lambda^\dagger c_\lambda$. We note also that if we introduce the variables $q = c^+ + c$ and $p = \frac{1}{i}(c^+ - c)$, the problem is transformed into the harmonic oscillator problem.

It is easily shown that the energy operator is of the form

$$H = \frac{1}{2} \sum_{\lambda, k} \hbar \omega (c_\lambda c_\lambda^\dagger + c_\lambda^\dagger c_\lambda). \quad (15.25)$$

The expression, unlike the classical one (15.20), contains longitudinal and scalar oscillations corresponding to $\lambda = 3, 4$, since in deriving (15.25) we did not make use of the subsidiary condition (15.6). We have seen above that when this subsidiary condition is applied to the potential $\mathbf{E}$ and $\mathbf{F}$ in classical electrodynamics depend on the oscillations with $\lambda = 3, 4$.

The fact that the energy and momentum do not depend on the "longitudinal" and "scalar" oscillations should, of course, remain true also in quantum electrodynamics, but the subsidiary condition in its usual classical form is incompatible with the quantum conditions. This can be seen directly by comparing (15.16) and (15.23), since in the quantum conditions (15.23) the variables c_3 and c_4 must be considered independent, whereas according to (15.16) $c_3 = -ic_4$.

We can avoid this difficulty by restricting the class of permissible wave functions Φ which describe the states of the field of the generalized quantum mechanical system.²⁾

In particular, we shall consider permissible only such wave functions Φ which are annihilated by the positive-frequency part of the operator $\frac{\partial}{\partial x_\mu} A_\mu$, that is by that part which contains terms $\frac{e^{ikx}}{A_\mu}$; in other words, we shall not consider the operator $\frac{\partial}{\partial x_\mu} A_\mu$ identically equal to zero, but shall allow only such states Φ which satisfy the condition

$$\left(\frac{\partial A_\mu}{\partial x_\mu} \right)_+ \Phi = 0, \quad (15.26)$$

where $\left(\frac{\partial}{\partial x_\mu} A_\mu \right)_+$ is the positive frequency part of $\frac{\partial}{\partial x_\mu} A_\mu$. According to (15.16) and (15.22) this condition can be written

$$\left(\sum_k V^\omega (c_3 + ic_4) e^{ikx} \right) \Phi = 0,$$

whence it follows that the wave function Φ for all vectors k must satisfy the condition

$$(c_3 + ic_4) \Phi = 0. \quad (15.27)$$

From (15.27) we obtain also

²⁾ E. Fermi, Revs. Mod. Phys. 4, 125 (1932).

$$\Phi^*(c_3^+ - ic_4^+) = 0 \quad (15.27')$$

and therefore, in addition to (15.26), Φ satisfies the condition

$$\Phi^* \left(\frac{\partial A_\mu}{\partial x_\mu} \right) = 0, \quad (15.26')$$

where $\left(\frac{\partial A_\mu}{\partial x_\mu} \right)_-$ is the negative-frequency part of $\frac{\partial A_\mu}{\partial x_\mu}$, or that part containing terms e^{-ikx} .

Taking the scalar product of (15.26) with Φ^* , and that of (15.26') with Φ , and adding the expressions so obtained we arrive at

$$\left(\Phi, \frac{\partial A_\mu}{\partial x_\mu} \Phi \right) = \left\langle \frac{\partial A_\mu}{\partial x_\mu} \right\rangle = 0, \quad (15.28)$$

where $\langle L \rangle = (\Phi, L\Phi)$ is the expectation value of the operator L in the state Φ .

We see that in quantum electrodynamics it is not the operator $\frac{\partial A_\mu}{\partial x_\mu}$ which vanishes, but its expectation value in the states Φ .

From (15.27) and (15.27') it follows that

$$(\Phi, (c_3^+ - ic_4^+)(c_3 + ic_4)\Phi) + (\Phi, (c_3^+ + ic_4^+)(c_3 - ic_4)\Phi) = 0,$$

or

$$(\Phi, (c_3^+ c_3 + c_4^+ c_4)\Phi) = \langle c_3^+ c_3 \rangle + \langle c_4^+ c_4 \rangle = 0. \quad (15.29)$$

We note that it does not follow from this that $N_3 = N_4 = 0$. On the contrary, as we shall see later, there exists no state of the field for which $N_3 = N_4 = 0$.

Let us now determine the eigenvalues of the energy operator (15.25). Noting that $c_\lambda c_\lambda^+ = c_\lambda^+ c_\lambda + 1$ and making use of (15.29), we find that the eigenvalues of the energy are

$$\mathcal{E}_{N_1, N_2, \dots} = \sum_{\lambda=1,2} \hbar \omega N_\lambda + \sum_{\lambda} \frac{1}{2} \hbar \omega. \quad (15.30)$$

We see that if we ignore the divergent sum $\sum_{\lambda} \frac{1}{2} \hbar \omega$, the electromagnetic field energy is actually given as the sum of the energies of separate photons. The integer $N_\lambda = N_{k\lambda}$ gives the number of photons with wave vector k , polarization λ ($\lambda = 1, 2$), and energy $\hbar \omega$.

The energy of the field is given only by the transverse oscillations ($\lambda = 1, 2$). Oscillations with $\lambda = 3, 4$ and the numbers of longitudinal and scalar "photons" N_3 and N_4 corresponding to them do not enter into the expression for the energy.

Similarly we find the eigenvalues of the momentum operator, which is given by the expression

$$P = \frac{1}{c} \sum_{k, \lambda} \hbar k (c_k c_k^+ + c_k^+ c_k). \quad (15.31)$$

The eigenvalues of the momentum are

$$P_{N_1, N_2, \dots} = \sum_{k, \lambda=1,2} \hbar k N_\lambda + \sum_{k, \lambda} \frac{1}{2} \hbar k. \quad (15.32)$$

We see that if we ignore the sum $\sum_{k, \lambda} \frac{1}{2} \hbar k$, the momentum of the field is the sum of the momenta of the separate photons, and that as in classical electrodynamics, the momentum depends only on the transverse oscillations of the field ($\lambda = 1, 2$).

Equation (15.30) shows that the energy of the field does not vanish when all photon occupation numbers N_λ ($\lambda = 1, 2$) are zero. The state with $N_\lambda = 0$ ($\lambda = 1, 2$) is the vacuum state of the electromagnetic field. We may, therefore, say that the quantization rules (15.23) lead to an infinite vacuum energy (this energy is called the zero-point energy).

One might think that the zero-point energy plays no role and can simply be dropped, since it does not enter into the differences of the energy eigenvalues, which are all that are of importance in energy transfer. This conclusion, however, would be incorrect, since as we shall see later (see Chapter VIII), the vacuum oscillations are of importance in many effects having to do with the interaction of charged particles with the electromagnetic field.

In Section 17 we shall see that the vacuum must be defined in the same way also for the electron-positron field, and that the quantum rules for the electron-positron field lead to infinite energy and infinite charge of the electron-positron vacuum. Just as the zero-point oscillations of the electromagnetic field cannot be ignored, neither can those of the electron-positron field, since their effect is felt in many phenomena related to interactions between fields.

The interaction of fields with zero-point vacuum oscillations leads to fundamental difficulties involving divergent expressions for the energies and probabilities of various interaction processes between electrons and the electromagnetic field. Nevertheless, as we shall see later (Sections 26, 27), general rules can be formulated for uniquely separating finite and physically meaningful quantities out of the divergent expressions.

The fact that the vacuum possesses physical properties and cannot simply be considered "empty" space is an extremely important result of quantum electrodynamics. Effects related to the interaction of electrons with the electromagnetic vacuum and of the electromagnetic field with the electron-positron vacuum will be given special consideration in Chapter VIII.

Let us now sum up. It can be said that quantization of the electromagnetic field reduces to considering the potential as an operator of the form.⁴⁾

⁴⁾ We assume here and in the future that $\hbar = c = 1$.

$$A_\mu = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}, \lambda} \frac{1}{\sqrt{2\omega}} (c_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + c_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}}) \epsilon_{\mu\lambda}, \quad (15.33)$$

where the operators c_λ and $c_\lambda^\dagger$ satisfy the quantum conditions (15.23).

The operator A_μ , just as the tensor field operator $F_{\mu\nu} = \frac{\partial}{\partial x_\mu} A_\nu - \frac{\partial}{\partial x_\nu} A_\mu$, acts on the wave function which describes the state of the field and depends on the number of photons $\mathbf{N}_\lambda$. Only those wave functions Φ are admissible which satisfy condition (15.27).

As a function of space and time, the operator A_μ satisfies D'Alembert's equation

$$\square A_\mu = 0 \quad (\mu = 1, 2, 3, 4). \quad (15.34)$$

This equation, together with the quantum conditions (15.23) and the subsidiary condition (15.27), is the complete set of equations of quantum electrodynamics for the free electromagnetic field.

From these equations we can now obtain Maxwell's equations, which can be written

$$\left. \begin{aligned} \frac{\partial}{\partial x_\sigma} F_{\mu\nu} + \frac{\partial}{\partial x_\nu} F_{\sigma\mu} + \frac{\partial}{\partial x_\mu} F_{\nu\sigma} &= 0, \\ (\Phi, \frac{\partial}{\partial x_\mu} F_{\mu\nu} \Phi) &= 0. \end{aligned} \right\} \quad (15.35)$$

We see that the operator $\frac{\partial}{\partial x_\sigma} F_{\mu\nu} + \frac{\partial}{\partial x_\nu} F_{\sigma\mu} + \frac{\partial}{\partial x_\mu} F_{\nu\sigma}$ as the corresponding quantity in classical electrodynamics, vanishes. As for $\frac{\partial}{\partial x_\mu} F_{\mu\nu}$, the situation is not exactly the same as in classical electrodynamics, where $\frac{\partial}{\partial x_\mu} F_{\mu\nu} = 0$; in quantum electrodynamics only the expectation value of the operator $\frac{\partial}{\partial x_\mu} F_{\mu\nu}$ in the state Φ vanishes. This is related to the fact that in quantum electrodynamics the subsidiary condition is not $\frac{\partial}{\partial x_\mu} A_\mu = 0$, but is $(\Phi, \frac{\partial}{\partial x_\mu} A_\mu \Phi) = 0$; it is this last equation together with (15.34) which leads to the second of Equations (15.35).

By taking the expectation value of the first of Equations (15.35) for a state Φ , we obtain the classical Maxwell's equations for the expectation value of the operator $F_{\mu\nu}$, namely

$$\left. \begin{aligned} \frac{\partial}{\partial x_\sigma} \langle F_{\mu\nu} \rangle + \frac{\partial}{\partial x_\nu} \langle F_{\sigma\mu} \rangle + \frac{\partial}{\partial x_\mu} \langle F_{\nu\sigma} \rangle &= 0, \\ \frac{\partial}{\partial x_\mu} \langle F_{\mu\nu} \rangle &= 0, \end{aligned} \right\} \quad (15.36)$$

where $\langle F_{\mu\nu} \rangle = (\Phi, F_{\mu\nu} \Phi)$ (the quantum mechanical average is taken over the variables on which Φ depends, i.e., on the photon occupation numbers, and, therefore, $\langle F_{\mu\nu} \rangle$ is a function of space and time).

3. Definition of the Vacuum for the Electromagnetic Field. The Use of the Indefinite Metric.

We have seen that the vacuum of the electromagnetic field is that state in which no transverse photons are present, i.e., in which the $\mathbf{N}_\lambda$ ($\lambda = 1, 2$) vanish. As for the longitudinal and scalar "photons," unlike the transverse ones, they are present in the vacuum state. Indeed, let us consider that part of the wave function $\Phi(\mathbf{N}_\lambda, \mathbf{N}_\lambda)$, which describes the longitudinal and scalar oscillations. This function can be written in the form

$$\Phi(\mathbf{N}_\lambda, \mathbf{N}_\lambda) = \sum_{N_0^0, N_4^0} C(N_0^0, N_4^0) \Phi_{N_0^0 N_4^0}(\mathbf{N}_\lambda, \mathbf{N}_\lambda), \quad (15.37)$$

where $\Phi_{N_0^0 N_4^0}(\mathbf{N}_\lambda, \mathbf{N}_\lambda)$ describes a state with definite numbers of longitudinal ($\mathbf{N}_0^0$) and scalar ($\mathbf{N}_4^0$) "photons".³⁾ Inserting this expression into (15.27) and making use of (15.24), we arrive at the relation

$$\sum_{N_0^0, N_4^0} C(N_0^0, N_4^0) \{ \sqrt{N_0^0} \Phi_{N_0^0-1, N_4^0}(\mathbf{N}_\lambda, \mathbf{N}_\lambda) - \sqrt{N_4^0+1} \Phi_{N_0^0, N_4^0+1}(\mathbf{N}_\lambda, \mathbf{N}_\lambda) \} = 0.$$

Performing the substitution $N_0^0 - 1 \rightarrow N_0^0$ in the first term, and $N_4^0 + 1 \rightarrow N_4^0$ in the second, this relation can be written

$$\sum_{N_0^0 > 0, N_4^0 > 0} C(N_0^0+1, N_4^0) \sqrt{N_0^0+1} \Phi_{N_0^0+1, N_4^0}(\mathbf{N}_\lambda, \mathbf{N}_\lambda) - \sum_{N_0^0 > 0, N_4^0 > 1} C(N_0^0, N_4^0-1) \sqrt{N_4^0-1} \Phi_{N_0^0, N_4^0-1}(\mathbf{N}_\lambda, \mathbf{N}_\lambda) = 0,$$

from which it follows that the function C satisfies the difference equation

$$C(N_0^0+1, N_4^0) \sqrt{N_0^0+1} = C(N_0^0, N_4^0-1) \sqrt{N_4^0-1}, \\ N_0^0 = 0, 1, \dots; N_4^0 = 1, 2, \dots$$

and the "boundary" condition

$$C(N_0^0, 0) = 0, \quad N_0^0 = 1, 2, \dots$$

³⁾ $\Phi_{N_0^0 N_4^0}(\mathbf{N}_\lambda, \mathbf{N}_\lambda) = \delta_{N_0^0 N_4^0} \delta_{N_0^0 N_4^0}$.

Replacing $\underline{C}$ by a new variable $\underline{B}$, defined by

$$C(N_b, N_d) = B(N_b, N_d) \sqrt{\frac{N_d}{N_b}},$$

we obtain

$$B(N_b + 1, N_d) = B(N_b, N_d - 1) = B(N_b - 1, N_d - 2) = \dots$$

In other words, B depends only on the difference $N_d - N_b$. Making use of the "boundary" condition, we find that the desired function $\underline{C}$ is

$$C(N_b, N_d) = \begin{cases} 0, & N_d < N_b, \\ \sqrt{\frac{N_d}{N_b}} f(N_d - N_b), & N_d \geq N_b, \end{cases} \quad (15.38)$$

where $f(\underline{M})$ is an arbitrary function of the nonnegative integer $\underline{M}$.

We see that various states are possible for the field of longitudinal and scalar oscillations, although there exists no state in which the number of longitudinal and scalar "photons" vanishes.

From (15.38) it follows that the wave function $\Phi(N_b, N_d)$ cannot be normalized, since $\sum_{N_b, N_d} |C(N_b, N_d)|^2 = \infty$, and thus, the usual definition of expectation values in the state $\Phi(N_b, N_d)$, does not, strictly speaking, have meaning.

The longitudinal and scalar "photons" do not enter into the expressions for the energy and momentum, and have no physical meaning for the free electromagnetic field. Their occurrence in the theory, however, presents formal difficulties. We shall present a method which can be used to avoid these difficulties. To do this we must change the quantization conditions for the scalar potential and use a definition which is different from the usual one for the expectation value of an operator.

According to the usual definition, the expectation value of an operator L in a state Φ is

$$\langle L \rangle = (\Phi, L\Phi) = \int \Phi^* L \Phi dq. \quad (15.39)$$

Here Φ^* is the complex conjugate of Φ , and dq is the product of the differentials of the variables on which Φ depends; it is assumed here that Φ is normalized according to

$$(\Phi, \Phi) = \int \Phi^* \Phi dq = 1. \quad (15.39')$$

We can, however, use a different definition of the expectation value of an operator L in a state Φ , namely¹⁾

$$\langle L \rangle = (\Phi, L\Phi) = \int \Phi^* L \Phi dq, \quad (15.40)$$

where

$$\Phi^* = \Phi^\dagger \eta,$$

and η is in general an arbitrary Hermitian operator. The normalization condition can now be written

$$(\Phi, \Phi) = \int \Phi^* \Phi dq = \pm 1. \quad (15.40')$$

We see that the norm of the wave function corresponding to this definition of the expectation value may be negative (an indefinite metric in Hilbert space).

The expectation value of an operator corresponding to any physical quantity should obviously be real. We shall prove that the expectation value of L is real if it satisfies the equation

$$L = L^*, \quad (15.41)$$

where

$$L^* = \eta^{-1} L^\dagger \eta \quad (15.42)$$

and $L^\dagger$ is the Hermitian conjugate of L . For simplicity let us consider the case of a discrete variable. In this case

$$\langle L \rangle = \sum_{nm} \Phi_n^* \eta_{nm} L_{mi} \Phi_i.$$

Since η is Hermitian by definition, the complex conjugate of $\langle L \rangle$ can be written

$$\langle L \rangle^* = \sum_{nm} \Phi_i^* (L^*)_{im} \eta_{mn} \Phi_n = \sum_{nm} \Phi_i^* \eta_{im} (L^*)_{nn} \Phi_n = \langle L^* \rangle.$$

¹⁾W. Pauli, Revs. Mod. Phys. 15, 175 (1943).

Thus, if (15.41) is satisfied, then $\langle L \rangle^* = \langle L \rangle$, and $\langle L \rangle$ is real.

The operator $L^* = \eta^{-1} L^* \eta$ shall be called the conjugate of L , and an operator satisfying Equation (15.41) shall be called self-conjugate.

When using the indefinite metric in Hilbert space, that is when the norm of a vector is defined not as

$\int \Psi^* \Psi d\mathbf{q}$, but as $\int \Psi^* \eta \Psi d\mathbf{q}$, operators which are self-conjugate in the sense of (15.41) play the same role as Hermitian operators in the usual scheme in which (15.39') is used as the definition of the norm.

The general definition of the expectation value in the form of (15.40), based on the use of the indefinite metric, can be employed in quantizing the electromagnetic field. In this way we avoid the difficulties related to the longitudinal and scalar "photons", in particular we eliminate the above-mentioned problem of the impossibility of the usual definition of the expectation values in the state $\Phi(\underline{N}_s, \underline{N}_s)$ describing the longitudinal and scalar oscillations.

Quantization on the basis of the indefinite metric is performed as follows.¹⁾

We start by expanding the potential A_μ in plane waves in the form

$$\left. \begin{aligned} A_j(x) &= \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega}} (c_{\mathbf{k}j} e^{ikx} + c_{\mathbf{k}j}^* e^{-ikx}) \epsilon_{j\mathbf{k}}, \quad j = 1, 2, 3, \\ A_0(x) &= \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega}} (c_{\mathbf{k}0} e^{ikx} + c_{\mathbf{k}0}^* e^{-ikx}) \epsilon_{0\mathbf{k}} \end{aligned} \right\} \quad (15.42)$$

and consider the operators $c_{\mathbf{k}\lambda} = c_{\mathbf{k}\lambda}$ and $c_{\mathbf{k}\lambda}^* = c_{\mathbf{k}\lambda}^*$ to be conjugate in the sense of (15.42), and to satisfy the quantum conditions

$$[c_j, c_j^*] = 1, [c_0, c_0^*] = -1. \quad (15.44)$$

The Hermitian operator η entering into (15.42) is defined by the conditions

$$\left. \begin{aligned} c_j^* &= c_j^*, \quad j = 1, 2, 3, \\ c_0^* &= -c_0^*, \end{aligned} \right\} \quad (15.45)$$

which can be written

$$\left. \begin{aligned} c_j^* \eta &= \eta c_j^*, \\ c_0^* \eta &= -\eta c_0^*. \end{aligned} \right\} \quad (15.45')$$

We note that conditions (15.44) differ from (15.23) and (15.23') in that they contain the symbol $*$ instead of $\dagger$.

From (15.44) and (15.45') it follows that the eigenvalues of the operators $c_j^* c_j = c_j^* c_j$ ($j = 1, 2, 3$) are $N_j = 0, 1, 2, \dots$, and the eigenvalues of $c_0^* c_0 = -c_0^* c_0$ are $-N_0 = 0, -1, -2, \dots$; N_j and N_0 are the numbers of "transverse" photons, and N_s and N_s are the numbers of longitudinal and scalar "photons".

We shall assume that the operator η is diagonal in the representation in which $c_j^* c_j$ and $c_0^* c_0$ are diagonal. It follows from (15.45') that its diagonal elements can be written

$$(N_1, N_2, N_3, N_0 | \eta | N_1, N_2, N_3, N_0) = \eta_{N_1 N_2 N_3 N_0} \eta_{N_1 N_2 N_3 N_0},$$

where the η_{N_j} satisfy the conditions

$$\left. \begin{aligned} \eta_{N_j+1} &= \eta_{N_j}, \quad j = 1, 2, 3, \\ \eta_{N_0+1} &= -\eta_{N_0}. \end{aligned} \right\}$$

These conditions show that the operator η can be considered the unit operator with respect to the variables N_j ($j = 1, 2, 3$), and to act as $(-1)^{N_0}$ with respect to the variable N_0 ; in other words, the matrix elements of η^{-1} are

$$(N_1, N_2, N_3, N_0 | \eta^{-1} | N_1', N_2', N_3', N_0') = (-1)^{N_0} \delta_{N_1 N_1'} \delta_{N_2 N_2'} \delta_{N_3 N_3'} \delta_{N_0 N_0'}. \quad (15.46)$$

Let us use $\Phi_j(N_j)$ ($j = 1, 2, 3, 0$) to denote the eigenfunction of a state containing N_j photons of the j -th kind. Then from (15.46') and (15.46) we find that these functions should be normalized according to

$$\left. \begin{aligned} \langle \Phi_j(N_j), \Phi_j(N_j') | &= \delta_{N_j N_j'}, \quad j = 1, 2, 3, \\ \langle \Phi_0(N_0), \Phi_0(N_0') | &= (-1)^{N_0} \delta_{N_0 N_0'}. \end{aligned} \right\} \quad (15.47)$$

Using (15.44) and (15.45) it can be shown that the application of the operators c_λ and c_λ^* to Φ_j and Φ_0 gives the following results:

$$\left. \begin{aligned} c_j \Phi_j(N_j) &= \sqrt{N_j} \Phi_j(N_j - 1), \\ c_j^* \Phi_j(N_j) &= \sqrt{N_j + 1} \Phi_j(N_j + 1), \quad j = 1, 2, 3, \\ c_0 \Phi_0(N_0) &= -\sqrt{N_0} \Phi_0(N_0 - 1), \\ c_0^* \Phi_0(N_0) &= \sqrt{N_0 + 1} \Phi_0(N_0 + 1). \end{aligned} \right\} \quad (15.48)$$

¹⁾ S. Gupta, Proc. Phys. Soc. (London) 63A, 681 (1950).

We see that the operators c_j ($j = 1, 2, 3$) and c_s are photon absorption operators, and that c_j^* and c_s^* are photon emission operators. We note that this is not the same as the result obtained when quantizing on the basis of (15.23), in which case the operator c_s was a scalar "photon" emission operator; in this case c_s , like the c_j , is an absorption operator.

Let us now establish the form of the permissible wave functions Φ of the electromagnetic field. The subsidiary condition (15.26) remains the same, and now

$$\left(\frac{\partial A_k}{\partial x_k}\right)_+ \Phi = 0. \quad (15.49)$$

Equation (15.26'), however, must now be written (15.49').

$$\Phi + \left(\frac{\partial A_k}{\partial x_k}\right)_- = 0. \quad (15.49')$$

From (15.49) we have

$$(c_0 - c_0^*) \Phi = 0. \quad (15.50)$$

We shall attempt to find Φ in the form

$$\Phi = \sum_{N_1, N_2, N_3, N_s} B_{N_1, N_2, N_3, N_s} \Phi(N_1, N_2, N_3, N_s). \quad (15.51)$$

where $\Phi(N_1, N_2, N_3, N_s)$ is a normalized wave function describing the state of the field in which there are N_1 and N_2 transverse, N_3 longitudinal, and N_s scalar photons, and the B_{N_1, N_2, N_3, N_s} are numerical coefficients.

Conditions (15.50) and (15.48) give

$$V \bar{N}_s B_{N_1, N_2, N_3, N_s-1, N_s} + V \bar{N}_0 B_{N_1, N_2, N_3, N_s, N_s-1} = 0 \quad (15.52)$$

(those B that have negative indices vanish), from which it follows that the following states of the free electromagnetic field are possible:

$$\left. \begin{aligned} \Phi_0 &= \Phi(N_1, N_2, 0, 0), \\ \Phi_1 &= \Phi(N_1, N_2, 1, 0) - \Phi(N_1, N_2, 0, 1), \\ \Phi_2 &= \Phi(N_1, N_2, 2, 0) - \sqrt{2} \Phi(N_1, N_2, 1, 1) + \Phi(N_1, N_2, 0, 2), \\ &\dots \end{aligned} \right\} \quad (15.53)$$

where N_1 and N_2 may take on the values 0, 1, 2, ...

We see that when quantizing according to (15.44) and using the indefinite metric for the definition of the expectation value of an operator, it is possible to obtain a state in which no longitudinal scalar "photons" are present (the state Φ_0); this is a different result than that obtained on the basis of the quantum conditions (15.23).

An arbitrary superposition of the states given by (15.53) results in a general state of the electromagnetic field, namely

$$\Phi = \Phi_0 + a_1 \Phi_1 + a_2 \Phi_2 + \dots \quad (15.54)$$

where $a_1, a_2, \dots$ are arbitrary numbers.

As opposed to (15.37), this function Φ is normalizable. Indeed, using (15.47) it is easily shown that

$$(\Phi, \Phi) = (\Phi_0, \Phi_0) \quad (15.55)$$

irrespective of $a_1, a_2, \dots$

Thus, the norm of Φ depends only on the state in which longitudinal and scalar "photons" do not exist. Noting that the energy operator in terms of the variables c_j and c_j^* is of the form

$$H = \sum_k \omega (c_1^* c_1 + c_2^* c_2 + c_3^* c_3 - c_s^* c_s)$$

(we have dropped the zero-point energy), it can easily be shown that the states $\Phi_1, \Phi_2, \dots$ give no contribution to the expectation value of the energy which, as might have been expected, is determined only by the transverse oscillations. We shall thus exclude all consideration of the states $\Phi_1, \Phi_2, \dots$ and consider the free electromagnetic field to contain only transverse photons, the number of longitudinal and scalar "photons" being zero:

$$N_3 = N_0 = 0. \quad (15.55')$$

4. Wave Function of a System of Photons in Momentum Space and Second Quantization.

We shall show how the quantum conditions for the electromagnetic field are related to the description of a system of photons by a momentum-space wave function.

A set of N photons can be described by a wave function $f(k_1 \alpha_1, k_2 \alpha_2, \dots, k_N \alpha_N)$ which depends on the photon momenta k_i and the variables α_i which determine the polarizations (each variable α_i can take on three values given by $\alpha_i = 1, 2, 3$). The wave function is symmetric with respect to interchange of particles,

$$f(\dots, k_i \alpha_i, \dots, k_j \alpha_j, \dots) = f(\dots, k_j \alpha_j, \dots, k_i \alpha_i, \dots)$$

and satisfies the transversality conditions (see Section 6)

$$\sum_{\alpha_i=1,2,\dots} f(\dots, k_i \alpha_i, \dots)(k_i)_{\alpha_i} = 0, \quad i=1, 2, \dots, N_f$$

where the summation is taken over the three values of the variable α_i (the number of the transversality conditions is equal to the number of photons).

If the state of each photon is characterized by the set of quantum numbers $\underline{i}$ (for instance, $\underline{p}$ or $\underline{p}, \underline{m}, \underline{\lambda}$), then the state of the system is completely determined by the "occupation numbers" $N_{\underline{i}}$, that is by the number of photons $N_{\underline{i}}$ in each state characterized by $\underline{i}$. The wave function of a system with a definite distribution of the occupation numbers can be expressed in terms of the individual photon wave functions $f_{\underline{i}}(k, \alpha)$:

$$f_{N_{\underline{i}} N_{\underline{j}} \dots}(k_1 \alpha_1, \dots, k_N \alpha_N) = \left(\frac{N_{\underline{i}}! N_{\underline{j}}! \dots}{N!} \right)^{1/2} \sum_{\substack{N = \sum_i N_{\underline{i}}}} f_{\underline{i}}(k_1 \alpha_1) f_{\underline{j}}(k_2 \alpha_2) \dots,$$

where the summation is taken over all permutations of the indices $\underline{i}_1$. The function f is normalized according to

$$\sum_{\alpha_1, \alpha_2, \dots} \int |f_{N_{\underline{i}} \dots}(k_1 \alpha_1, \dots, k_N \alpha_N)|^2 dk_1 dk_2 \dots dk_N = 1.$$

An arbitrary state of a system of N photons can be represented in the form of a superposition of states with definite occupation numbers,

$$f_N(k_1 \alpha_1, \dots) = \sum_{N_{\underline{i}} \dots N_{\underline{j}}} \Phi(N_{\underline{i}}, \dots, N_{\underline{j}}) f_{N_{\underline{i}} \dots N_{\underline{j}}}(k_1 \alpha_1, k_2 \alpha_2, \dots). \quad (15.56)$$

In the interaction between photons and charges, the number of photons may vary, since photons may be emitted and absorbed. Therefore, in general it is necessary to consider states with a variable number of particles. Such states are described by a collection of wave functions $f_{\underline{i}}$, which can be written as a column vector

$$\begin{pmatrix} a_0 \\ a_1 f_1(k_1 \alpha_1) \\ a_2 f_2(k_1 \alpha_1, k_2 \alpha_2) \\ \dots \\ a_N f_N(k_1 \alpha_1, \dots, k_N \alpha_N) \end{pmatrix}, \quad (15.57)$$

where the square modulus of $a_{\underline{i}}$ determines the probability of finding $\underline{i}$ photons.¹⁾

The description of the electromagnetic field with the aid of the field potential operators, as we did in the previous sections, is equivalent to transition from (15.57) to another representation in which the wave functions are the quantities $\Phi(N_{\underline{i}}, \dots, N_{\underline{j}}, \dots)$ in (15.56). The functions $\Phi(N_{\underline{i}}, \dots, N_{\underline{j}}, \dots)$ completely determine $f_N(\dots, k_N \alpha_N, \dots)$. A function $\Phi(N_{\underline{i}}, \dots, N_{\underline{j}}, \dots)$ is nothing other than our previously introduced wave function for describing the electromagnetic field as a quantum mechanical system. This method of description is called second quantization.

The transition to the occupation-number representation, that is, the transformation to the new representation for the energy, momentum, and angular momentum operators of the field, is performed, as is well known,²⁾ by introducing the emission and absorption (or creation and annihilation) operators for a particle in the state $\underline{i}$. These operators, which we shall denote by $c_{\underline{i}}^+$ and $c_{\underline{i}}$, are defined by the relations

$$\left. \begin{aligned} c_{\underline{i}} \Phi(\dots, N_{\underline{i}}, \dots) &= \sqrt{N_{\underline{i}}} \Phi(\dots, N_{\underline{i}} - 1, \dots), \\ c_{\underline{i}}^+ \Phi(\dots, N_{\underline{i}}, \dots) &= \sqrt{N_{\underline{i}} + 1} \Phi(\dots, N_{\underline{i}} + 1, \dots) \end{aligned} \right\} \quad (15.58)$$

and are identical with the operators introduced by (15.23).

Let us construct the operators

$$\left. \begin{aligned} F(k, \alpha) &= \sum_{\underline{i}} c_{\underline{i}} f_{\underline{i}}(k, \alpha), \\ F^+(k, \alpha) &= \sum_{\underline{i}} c_{\underline{i}}^+ f_{\underline{i}}^*(k, \alpha), \end{aligned} \right\} \quad (15.59)$$

where $f_{\underline{i}}(k, \alpha)$ is the wave function of the photon in the state $\underline{i}$. The operator F is called the quantized wave function of the photon. The operators F and F^+ satisfy the commutation rules

$$F(k, \alpha) F^+(k', \alpha') - F^+(k', \alpha') F(k, \alpha) = \delta_{kk'} \delta_{\alpha\alpha'}, \quad (15.59')$$

which follow directly from (15.23) and (15.59). The operators of the potentials $A_{\mu}(x)$ can be expressed in terms of F and F^+ , just as the electromagnetic field potentials corresponding to a single photon are expressed in terms of its wave function $f(k, \alpha)$ (see Section 5).

5. Method of Fock Functionals.

Thus, a system of photons (in general the number of photons is not fixed) can be described either by a wave function $\Phi(\underline{n}_1, \underline{n}_2, \dots)$ in the space of occupation numbers, or by a set of wave functions $f_{\underline{i}}(k_1 \alpha_1, k_2 \alpha_2, \dots, k_N \alpha_N)$ as in (15.57) in momentum (and spin) space, each function corresponding to a definite number of particles

¹⁾ This method was developed by V. A. Fock [Z. Phys. 75, 622 (1932)].

²⁾ See, for instance, L. Landau and E. Lifshits, Quantum Mechanics (State Tech. Press, 1948); D. Blokhintsev, Elements of Quantum Mechanics (State Tech. Press, 1948).

$\underline{n}$. We shall now give a brief description of a method, due to V. A. Fock,¹⁾ for describing a system of photons, in which a certain functional plays the role of the wave function (this method is also applicable to arbitrary Bose particles).

Let us denote the momentum and spin of the photon by ξ . Correspondingly, we shall write the wave function of a system of $\underline{n}$ photons in the form $f_{\underline{n}}(\xi_1, \xi_2, \dots, \xi_n)$. Let $\varphi(\xi)$ be some function of ξ . Each wave function $f_{\underline{n}}$ can be associated in a one-to-one way with the functional

$$\Omega_n = \frac{1}{\sqrt{n!}} \int f_n(\xi_1, \xi_2, \dots, \xi_n) \varphi(\xi_1) \varphi(\xi_2) \dots \varphi(\xi_n) d\xi_1 \dots d\xi_n \quad (15.60)$$

(the integration involves also summation over the spin variable).

As an illustration, let us consider $\underline{n}$ photons in an eigenstate of ξ . In occupation-number space, this system is described by the wave function

$$\Phi(n_1, n_2, \dots, n_i, \dots) = \delta_{n_1} \delta_{n_2} \dots \delta_{n_i} \dots$$

In momentum space the wave function can be written

$$f_n(\xi_1, \dots, \xi_n) = \delta(\xi_1 - \xi) \delta(\xi_2 - \xi) \dots \delta(\xi_n - \xi).$$

This wave function, according to (15.60), corresponds to the functional

$$\Omega_n = \frac{1}{\sqrt{n!}} \varphi^n(\xi). \quad (15.61)$$

In general, a system whose number of photons is not fixed corresponds to the functional

$$\Omega = \sum_{n=0}^{\infty} a_n \Omega_n. \quad (15.62)$$

The scalar product of two functionals Ω and Ω' is defined

$$(\Omega, \Omega') = \sum_n a'_n a_n (\Omega_n, \Omega'_n) = \sum_n a'_n a_n \int f'_n(\xi_1, \dots, \xi_n) f_n(\xi_1, \dots, \xi_n) d\xi_1 \dots d\xi_n.$$

¹⁾ V.A. Fock, Sov. Phys. 6, 425 (1934).

If we know some operator U in the occupation-number representation, we can find its form U' in the representation in which the functional Ω plays the role of the wave function. It is entirely determined by the relation

$$(\Phi', U \Phi) = (\Omega', U' \Omega).$$

We shall show that the wave function operator $F^+(\xi)$ reduces to the operator of multiplication by $\varphi(\xi)$. This is easiest to see by using (15.61). The operator $F^+(\xi)$ corresponds to the creation of a particle at the point ξ , that is

$$F^+(\xi) \Phi(n_1, \dots, n_i, \dots) = \sqrt{n_i + 1} \Phi',$$

where

$$\Phi' = \delta_{n_1} \dots \delta_{n_i, n_i+1} \dots$$

On the other hand, the wave function Φ' corresponds to the functional

$$\Omega'_{n+1} = \frac{1}{\sqrt{(n+1)!}} \varphi^{n+1} = \frac{1}{\sqrt{n+1}} \varphi \Omega_n.$$

Thus, if we demand that

$$F^+(\xi) \Omega_n = \sqrt{n+1} \Omega'_{n+1}, \quad (15.63)$$

we obtain

$$F^+(\xi) \Omega = \varphi(\xi) \Omega. \quad (15.64)$$

Although this calculation has been performed for a special case, it is true also in general. Indeed, let us consider the product

$$\varphi(\xi) \Omega = \sum_n \frac{1}{\sqrt{n!}} \int f_n(\xi_1, \dots, \xi_n) \varphi(\xi_1) \dots \varphi(\xi_n) \varphi(\xi) d\xi_1 \dots d\xi_n.$$

If we introduce the symmetrized wave function of $(n+1)$ particles (an additional photon at the "point" ξ has been added to the n original photons)

$$f'_{n+1}(\xi_1, \dots, \xi_n, \xi_{n+1}) = \frac{1}{n+1} (f_n(\xi_1, \dots, \xi_n) \delta(\xi_{n+1} - \xi) + f_n(\xi_1, \dots, \xi_{n-1}, \xi_{n+1}) \delta(\xi_n - \xi) + \dots), \quad (15.65)$$

then $\varphi \Omega$ can be written

$$\varphi \Omega = \sum_n \frac{1}{\sqrt{n!}} \int f'_{n+1}(\xi_1, \dots, \xi_{n+1}) \varphi(\xi_1) \dots \varphi(\xi_{n+1}) d\xi_1 \dots d\xi_{n+1}$$

and introducing the functional Ω'_{n+1} corresponding to the wave function f'_{n+1} according to (15.60), we obtain

$$\varphi \Omega = \sum_n \sqrt{n+1} \Omega'_{n+1},$$

which verifies (15.63).

Let us now find the form of the operator F . This can be done by using the commutation relations (15.59):

$$F(\xi) F^+(\xi') = F^+(\xi') F(\xi) + \delta(\xi - \xi').$$

Since $F^+(\xi) \equiv \varphi(\xi)$, we have

$$F(\xi) \Omega = \frac{\delta \Omega}{\delta \varphi(\xi)}, \quad (15.66)$$

where $\frac{\delta}{\delta \varphi}$ is the variational derivative of the functional.¹⁾

¹⁾ This can be defined either by the expression that the variation $\delta \Omega = \int \frac{\delta \Omega}{\delta \varphi(\xi)} \delta \varphi(\xi) d\xi$, or by

$$\frac{\delta \Omega}{\delta \varphi(\xi)} = \lim_{\gamma \rightarrow 0} \frac{1}{\gamma} (\Omega[\varphi(\xi) + \gamma \delta(\xi' - \xi)] - \Omega[\varphi(\xi)]).$$

From the definition of the functional Ω_n , it is easy to see that

$$F(\xi) \Omega_n = \frac{\delta \Omega_n}{\delta \varphi(\xi)} = \sqrt{n} \Omega'_{n-1}, \quad (15.67)$$

where

$$\Omega'_{n-1} = \frac{1}{\sqrt{(n-1)!}} \int f'_{n-1}(\xi_1, \dots, \xi_{n-1}) \varphi(\xi_1) \dots \varphi(\xi_{n-1}) d\xi_1 \dots d\xi_{n-1}$$

is the functional corresponding to the wave function

$$f'_{n-1} = f_n(\xi_1, \dots, \xi_{n-1}, \xi), \quad (15.68)$$

In particular, in the example (15.61), $F \Omega_n = \frac{\sqrt{n}}{\sqrt{(n-1)!}} \varphi^n \Omega^{n-1}$. Equations (15.63) and (15.67) are equivalent to the corresponding equations (15.58) in the occupation-number representation.

If instead of the functionals Ω , we use the column function (15.57), it follows from (15.63) and (15.67) that

$$F^+(\xi) f_n(\xi_1, \dots, \xi_n) = \sqrt{n+1} f'_{n+1}(\xi_1, \dots, \xi_n, \xi), \\ F(\xi) f_n(\xi_1, \dots, \xi_n) = \sqrt{n} f'_{n-1}(\xi_1, \dots, \xi_{n-1}),$$

where f'_{n+1} is given by (15.65), and f'_{n-1} by (15.68). (The index corresponds to the position of the function in the column vector.) In more detail, this implies the following relations:

$$F(\xi) \begin{pmatrix} a_0 \\ a_1 f_1(\xi_1) \\ a_2 f_2(\xi_1, \xi_2) \\ a_3 f_3(\xi_1, \xi_2, \xi_3) \\ \vdots \end{pmatrix} = \begin{pmatrix} a_1 f_1(\xi) \\ a_2 \sqrt{2} f_2(\xi, \xi) \\ a_3 \sqrt{3} f_3(\xi, \xi, \xi) \\ \vdots \end{pmatrix}, \\ F^+(\xi) \begin{pmatrix} a_0 \\ a_1 f_1(\xi_1) \\ a_2 f_2(\xi_1, \xi_2) \\ a_3 f_3(\xi_1, \xi_2, \xi_3) \\ \vdots \end{pmatrix} = \begin{pmatrix} a_0 \delta(\xi_1 - \xi) \\ a_1 \sqrt{2} [f_1(\xi_1) \delta(\xi_2 - \xi) + f_1(\xi_2) \delta(\xi_1 - \xi)] \\ a_2 \sqrt{3} [f_2(\xi_1, \xi_2) \delta(\xi_3 - \xi) + \dots] \\ \vdots \end{pmatrix}. \quad (15.69)$$

§ 16. Commutators of the Electromagnetic Field. The singular Functions $\underline{D}$, $\underline{D}^{(1)}$, $\underline{D}^{(2)}$.

1. Quantum Conditions for the Potential.

Knowing the quantum conditions satisfied by c_λ and $c_\lambda^\dagger$, it is easy to find those satisfied by the operators A_μ . We shall show that these conditions reduce to the fact that the commutator bracket $[A_\mu(x), A_\nu(x')] = A_\mu(x)A_\nu(x') - A_\nu(x')A_\mu(x)$, sometimes called simply the commutator of $A_\mu(x)$ and $A_\nu(x')$, is a certain c-number.

In order to calculate $[A_\mu(x), A_\nu(x')]$, let us make use of the expansion (15.33), writing

$$[A_\mu(x), A_\nu(x')] = \frac{1}{2V} \sum_{\mathbf{k}, \mathbf{k}'} \frac{1}{\sqrt{\omega\omega'}} [(c_{\lambda\mu}c_{\lambda'\nu} - c_{\lambda'\mu}c_{\lambda\nu}) e^{i(kx + k'x')} + (c_{\lambda\mu}^\dagger c_{\lambda'\nu}^\dagger - c_{\lambda'\mu}^\dagger c_{\lambda\nu}^\dagger) e^{-i(kx + k'x')} + (c_{\lambda\mu}^\dagger c_{\lambda'\nu} - c_{\lambda'\mu}^\dagger c_{\lambda\nu}) e^{-i(kx - k'x')}] \epsilon_{\lambda\mu} \epsilon_{\lambda'\nu} \quad (16.1)$$

According to (15.23) and (15.15) this expression can be rewritten

$$[A_\mu(x), A_\nu(x')] = \frac{1}{2V} \sum_{\mathbf{k}} \frac{1}{\omega} [e^{ik(x-x')} - e^{-ik(x-x')}] \delta_{\mu\nu} \quad (16.2)$$

Going over from summation to integration (according to $(\sum_{\mathbf{k}} \dots = \int \dots \frac{V d^3k}{(2\pi)^3})$), we finally obtain

$$[A_\mu(x), A_\nu(x')] = -iD(x - x') \delta_{\mu\nu} \quad (16.3)$$

where¹⁾

$$D(x) = \frac{1}{(2\pi)^3} \int e^{ikx} \frac{\sin \omega t}{\omega} d^3k. \quad (16.4)$$

It is easy to see that the function $D(x)$ is invariant under Lorentz transformations. This follows directly from the fact that $D(x)$ can be written in the form

$$D(x) = \frac{i}{(2\pi)^3} \int e(k) e^{ikx} \delta(k^2) d^4k, \quad (16.4')$$

where $\epsilon(k) = \frac{k_0}{|k_0|}$ and $d^4k = d^3k dk_0$, $kx = kx - k_0t$.

Integrating (16.4) over the angle between $\mathbf{k}$ and $\mathbf{x}$, we arrive at

$$D(x) = \frac{1}{8\pi r} \int_0^\infty [e^{i\omega(r-t)} + e^{-i\omega(r-t)} - e^{i\omega(r+t)} - e^{-i\omega(r+t)}] d\omega = \frac{1}{4\pi r} [\delta(r-t) - \delta(r+t)], \quad (16.5)$$

where $r = |\mathbf{x}|$.

Thus, $\underline{D}(\underline{x})$ vanishes everywhere except on the light cone $\underline{x}^2 = \underline{x}^2 - \underline{x}^2 = 0$, and on the light cone it has a δ -function type of singularity.

It can be shown that $\underline{D}(\underline{x})$ is a singular solution of the equation

$$\square D(x) = 0, \quad (16.6)$$

with the boundary conditions

$$D(x, 0) = 0, \quad \left(\frac{\partial D}{\partial t}\right)_{t=0} = \delta(x). \quad (16.7)$$

With the aid of (16.3) we can find the commutation rules for the components of the field tensor, namely

$$[F_{\mu\nu}(x), F_{\mu'\nu'}(x')] = i \left\{ \delta_{\mu\nu'} \frac{\partial^2}{\partial x_\mu \partial x_{\nu'}} - \delta_{\mu\nu} \frac{\partial^2}{\partial x_{\mu'} \partial x_{\nu'}} - \delta_{\mu\nu'} \frac{\partial^2}{\partial x_{\mu'} \partial x_\nu} + \delta_{\mu\nu} \frac{\partial^2}{\partial x_\mu \partial x_{\nu'}} \right\} D(x - x'). \quad (16.8)$$

We see that the components of the field commute unless they are taken at space-time points lying on the same light cone, that is unless $(x - x')^2 = 0$.

The quantum conditions can be interpreted in the following well-known way. Let the operators $\underline{L}$ and $\underline{M}$ corresponding to variables $\underline{l}$ and $\underline{m}$ of some quantum mechanical system satisfy the commutation relations

$$[L, M] = C, \quad C \neq 0.$$

¹⁾ Jordan and Pauli, Z. Physik 47, 151 (1928).

We may then assert that there exists no state of the system which is a simultaneous eigenstate of both L and M ; a state of the system must be characterized by a set of values L and M , and the intervals ΔL and ΔM within which these values lie are related by an uncertainty principle

$$\Delta L \cdot \Delta M \geq |C|. \quad (16.9)$$

Equation (16.9) is the same kind of an uncertainty principle as that relating the coordinates and momenta in ordinary quantum mechanics.

We shall not here analyze the uncertainty relations which follow from (16.8) for the fields, but shall restrict ourselves to considerations of relations such as (16.9) which follow directly from the quantum conditions for c_λ and c_λ^+ ($\lambda = 1, 2$).¹⁾

Instead of c_λ and c_λ^+ let us introduce the new variables N_λ and ϕ_λ , given by

$$c_\lambda = e^{-i\phi_\lambda} \sqrt{N_\lambda}, \quad c_\lambda^+ = \sqrt{N_\lambda} e^{i\phi_\lambda}. \quad (16.10)$$

The eigenvalues of the operators N_λ are obviously the occupation numbers N_λ . Therefore, N_λ can be interpreted as the number-of-photons operator. The operator ϕ_λ also has a simple meaning: inserting (16.10) and (15.33) shows that ϕ_λ can be interpreted as the phase operator corresponding to the variables k, λ .

Let us establish the commutation relations between N_λ and ϕ_λ . Inserting (16.10) into (15.23), we obtain

$$N_\lambda e^{i\phi_\lambda} = e^{i\phi_\lambda} N_\lambda = e^{i\phi_\lambda}. \quad (16.11)$$

This commutation relation is satisfied if N_λ and ϕ_λ satisfy the commutation relation

$$N_\lambda \phi_\lambda - \phi_\lambda N_\lambda = -i. \quad (16.12)$$

Indeed, from (16.12) we obtain

$$N_\lambda \phi_\lambda^n - \phi_\lambda^n N_\lambda = -in \phi_\lambda^{n-1},$$

and if we write $e^{i\phi_\lambda} = \sum_{n=0}^{\infty} \frac{1}{n!} (i\phi_\lambda)^n$, we obtain (16.11).

¹⁾ See W. Heitler, The Quantum Theory of Radiation (State Tech. Press, 1940).

From the commutation relation (16.12), we arrive at the inequality

$$\Delta N_\lambda \Delta \phi_\lambda \geq 1. \quad (16.13)$$

The physical meaning of this is that the electromagnetic field cannot be characterized simultaneously by the number of various kinds of photons and definite phase relations.

If the phase difference between two waves has a well defined value, then only the total number of photons in these waves can be given. How these photons are distributed between these two waves, however, remains entirely undetermined.

The classical electromagnetic field is characterized by very large expectation values for the number of photons. In this case, even for small $\Delta \phi$ the expectation value $\langle N \rangle$ is much greater than ΔN ; this means, practically speaking, that the field is characterized both by definite phase relations and definite intensities (numbers of photons).

2. Expectation Value of the Operator $[A_\mu(x), A_\nu(x')]$ in the Vacuum State.

Later we shall have need of various products of the components of the potential operator $A_\mu(x)$, in particular, we will need an expression for the expectation value of $A_\mu(x) A_\nu(x') + A_\nu(x') A_\mu(x)$ in the vacuum state. In order to calculate this quantity we must know the expectation values of products of c_λ and c_λ^+ in the vacuum state, and we shall determine these here.

Since in the vacuum state the number of transverse photons is zero, we have

$$\langle c_\lambda^+ c_\lambda \rangle_0 = 0, \quad \langle c_\lambda c_\lambda^+ \rangle_0 = 1, \quad \lambda = 1, 2$$

($\langle L \rangle_0$ is the expectation value of the operator L in the vacuum state). The vacuum expectation values of $c_\lambda c_\lambda$ and $c_\lambda^+ c_\lambda^+$ are also zero, and therefore, for the vacuum state we may write

$$\langle c_\lambda c_\lambda' \rangle_0 = \delta_{\lambda\lambda'} \delta_{kk'}, \quad \langle c_\lambda^+ c_\lambda' \rangle_0 = 0, \quad \langle c_\lambda c_\lambda' \rangle_0 = 0, \quad \langle c_\lambda^+ c_\lambda'^+ \rangle_0 = 0. \quad (16.14)$$

These relations refer to the transverse oscillations. We shall now show²⁾ that (16.14) can be used also for the longitudinal and scalar oscillations, which then means that it is valid for all λ ($\lambda = 1, 2, 3, 4$). Let us first recall that the usual definition of expectation values in the state $\Phi(N_\mu, N_\mu')$ is not meaningful. We have shown in Section 15 that this difficulty can be avoided if we define expectation values of operators by using the indefinite metric in Hilbert space. We shall not, however, specialize the definition of expectation values in the state $\Phi(N_\mu, N_\mu')$ but shall assume that the expectation value of $c_\lambda^+ c_\lambda$ is somehow or other defined in the vacuum state. Let us denote this expectation value by $\langle c_\lambda^+ c_\lambda \rangle_0(k)$, where $\phi(k)$ is some given function, so that

$$\langle c_\lambda^+ c_\lambda \rangle_0 = \langle \phi_\lambda k \rangle_0 \phi(k). \quad (16.15)$$

²⁾ P. Dyson, Phys. Rev. 77, 420 (1950).

From (15.29) and (15.14) it follows in this case that

$$\langle c_i^+ c_i \rangle_0 = -\langle c_i^+ c_i \rangle_0 = (e_i k)^2 \varphi(k).$$

Further, from (15.27), (15.27'), and (15.14) we have

$$\langle c_i^+ c_i \rangle_0 = \langle c_i^+ c_i \rangle_0 = i \langle c_i^+ c_i \rangle_0 = i (e_i k)^2 \varphi(k) = (e_i k) (e_i k) \varphi(k).$$

All these relations, as well as (16.14), which are valid for $\lambda = 1, 2$, can be written

$$\left. \begin{aligned} \langle c_i^+ c_{i'} \rangle_0 &= (e_i k) (e_{i'} k) \varphi(k) \delta_{kk'}, & \langle c_i c_{i'} \rangle_0 &= 0, \\ \langle c_i^+ c_{i'}^+ \rangle_0 &= 0 & (\lambda, i' = 1, 2, 3, 4). \end{aligned} \right\} \quad (16.16)$$

From this and from (15.23) we obtain

$$\langle c_i c_{i'}^+ \rangle_0 = [\delta_{ii'} + (e_i k) (e_{i'} k) \varphi(k)] \delta_{kk'} \quad (\lambda, i' = 1, 2, 3, 4). \quad (16.16')$$

Defining the anticommutator bracket $\{c_i^+ c_{i'}\} = c_i^+ c_{i'} + c_{i'} c_i^+$, we can write

$$\left. \begin{aligned} \langle \{c_i^+ c_{i'}\} \rangle_0 &= \langle c_i c_{i'}^+ \rangle_0 = [\delta_{ii'} + 2\varphi(k) (e_i k) (e_{i'} k)] \delta_{kk'}, \\ \langle \{c_i c_{i'}\} \rangle_0 &= \langle \{c_i^+ c_{i'}^+\} \rangle_0 = 0. \end{aligned} \right\} \quad (16.17)$$

Let us now use these equations to determine the vacuum expectation value of the bracket

$$\{A_\mu(x), A_\nu(x')\} = A_\mu(x) A_\nu(x') + A_\nu(x') A_\mu(x).$$

Replacing $A_\mu(x)$ by the expansion (15.33), and using (16.17), we obtain

$$\langle \{A_\mu(x), A_\nu(x')\} \rangle_0 = \frac{1}{2V} \sum_{k, l, l'} \frac{1}{\omega} [\delta_{ll'} + 2\varphi(k) (e_l k) (e_{l'} k)] \times \\ \times (e^{ik(x-x')} + e^{-ik(x-x')}) \delta_{\mu\nu} e_{l'} e_{l'}.$$

Noting that $\delta_{\lambda\mu} \delta_{\lambda\sigma} = \delta_{\mu\sigma}$, this equation can be written

$$\langle \{A_\mu(x), A_\nu(x')\} \rangle_0 = \delta_{\mu\nu} \frac{1}{2V} \sum_k \frac{1}{\omega} (e^{ik(x-x')} + e^{-ik(x-x')}) + \\ + \frac{1}{V} \sum_k \frac{\varphi(k)}{\omega} (e^{ik(x-x')} + e^{-ik(x-x')}) \delta_{\mu\nu} k^2. \quad (16.18)$$

Let us introduce the functions

$$\left. \begin{aligned} D^{(1)}(x) &= \frac{1}{2V} \sum_k \frac{1}{\omega} (e^{ikx} + e^{-ikx}), \\ \chi(x) &= -\frac{1}{V} \sum_k \frac{\varphi(k)}{\omega} (e^{ikx} + e^{-ikx}). \end{aligned} \right\} \quad (16.18')$$

The second sum in (16.18) is obviously $\frac{\partial^2}{\partial x_\mu \partial x_\nu} \chi(x-x')$, so that (16.18) can be rewritten

$$\langle \{A_\mu(x), A_\nu(x')\} \rangle_0 = \delta_{\mu\nu} D^{(1)}(x-x') + \frac{\partial^2}{\partial x_\mu \partial x_\nu} \chi(x-x'). \quad (16.19)$$

As we shall see later (see Section 22), $\langle \{A_\mu(x), A_\nu(x')\} \rangle_0$ always enters the theory in an integral of the form

$$M = \int_{-\infty}^{+\infty} K_{\mu\nu}(x, x') \langle \{A_\mu(x), A_\nu(x')\} \rangle_0 d^4x d^4x', \quad (16.20)$$

where $K_{\mu\nu}(x, x')$ satisfies the conditions

$$\frac{\partial K_{\mu\nu}}{\partial x_\mu} = 0, \quad \frac{\partial K_{\mu\nu}}{\partial x_\nu} = 0, \quad (16.20')$$

and the integration is taken over all of four-space.

Integrating (16.20) by parts and using (16.20'), it is easy to see that

$$\int K_{\mu\nu}(x, x') \frac{\partial^2}{\partial x_\mu \partial x_\nu} \chi dx dx' = 0.$$

Therefore, the second term in (16.19) does not contribute to the integral $\mathcal{M}$, and since this is all that arises in the theory, we may consider the function χ as well, therefore, as the arbitrary function $\varphi(x)$ equal to zero. In other words, we may henceforth use relations (16.14) for all values of λ from $\lambda = 1$ to $\lambda = 4$.

Thus, we write

$$\langle [A_\mu(x), A_\nu(x')] \rangle_0 = D^{(1)}(x - x') \delta_{\mu\nu}. \quad (16.21)$$

Let us examine the properties of the function $D^{(1)}(x)$ in more detail. Using (16.18'), which defines $D^{(1)}(x)$, and going from summation over k to integration over $\frac{d^3k}{(2\pi)^3}$, we can write $D^{(1)}(x)$ in the form

$$D^{(1)}(x) = \frac{1}{(2\pi)^3} \int e^{ikx} \frac{\cos \omega t}{\omega} d^3k = \frac{1}{(2\pi)^3} \int e^{ikx_0} (k^2) d^4k \quad (16.22)$$

This last expression shows that $D^{(1)}(x)$ is a solution of the wave equation $\square D^{(1)} = 0$ and is invariant under Lorentz transformations. This function $D^{(1)}(x)$, just as $D(x)$, is a singular solution of the wave equation; unlike $D(x)$ which vanishes everywhere except on the light column $x_0 = 0$, the function $D^{(1)}(x)$ does not vanish for finite x_0 . It is easy to show that

$$D^{(1)}(x) = \frac{2}{(2\pi)^2} \frac{1}{x^2}, \quad x^2 = r^2 - t^2. \quad (16.22')$$

Indeed, integrating (16.22) over the angle between k and x , we obtain

$$\begin{aligned} D^{(1)}(x) &= \frac{1}{(2\pi)^3} \int \frac{\pi}{r} \int_0^\pi [e^{i\omega(r+t)} - e^{-i\omega(r-t)} + e^{i\omega(r-t)} - e^{-i\omega(r+t)}] d\omega = \\ &= \frac{1}{(2\pi)^3} \int \frac{\pi}{r} \left\{ \lim_{\epsilon \rightarrow 0} \int_0^\infty e^{-i\omega} [e^{i\omega(r+t)} - e^{-i\omega(r+t)}] d\omega + \right. \\ &\quad \left. + \lim_{\epsilon' \rightarrow 0} \int_0^\infty e^{-i'\omega} [e^{i\omega(r-t)} - e^{-i\omega(r-t)}] d\omega \right\} = \frac{2}{(2\pi)^2} \frac{1}{r^2 - t^2} \end{aligned}$$

where $\epsilon' = \epsilon \operatorname{sgn}(t - \tau)$ and $\epsilon > 0$.

3. Chronological Product of the Operators $A_\mu(x)$ and $A_\nu(x')$. The D^E Function.

Let us now introduce the so-called chronological product of the operators $A_\mu(x)$ and $A_\nu(x')$. This product is defined to mean that the operators A are chronologically ordered; that is, the operator on the right corresponds to the earliest time, and that on the left to the latest.

Let us calculate the vacuum expectation value of the chronological product of $A_\mu(x)$ and $A_\nu(x')$, denoted $P(A_\mu(x) A_\nu(x'))$.

We shall assume that $t > t'$. Then

$$P(A_\mu(x) A_\nu(x')) = \frac{1}{2V} \sum_{k_1, k'_1} \frac{1}{\sqrt{\omega\omega'}} \{ c_k c_{k'} e^{i(kx+k'x')} + \quad (16.23)$$

$+ c_k c_{k'}^* e^{i(kx-k'x')} + c_k^* c_{k'} e^{-i(kx-k'x')} + c_k^* c_{k'}^* e^{-i(kx+k'x')} \} \epsilon_{\mu\nu} \epsilon_{k_1 k'_1}$
[If $t < t'$, then the primed and unprimed variables should be interchanged everywhere in (16.23)]. Using (16.14) and the orthogonality of the $\epsilon_{\lambda\lambda'}$, we obtain

$$\langle P(A_\mu(x) A_\nu(x')) \rangle_0 = \frac{1}{2V} \sum_k \frac{1}{\omega} e^{i[kx - (x_0 - x'_0) - \omega(t - t')]} \delta_{\mu\nu}. \quad (16.24)$$

In this form this equation is obviously valid also for $t < t'$, since not $t - t'$, but $|t - t'|$ appears in the exponent. Replacing summation by integration over $\frac{d^3k}{(2\pi)^3}$, we can rewrite (16.24) in the form

$$\langle P(A_\mu(x) A_\nu(x')) \rangle_0 = \frac{1}{2} D^E(x - x') \delta_{\mu\nu}, \quad (16.25)$$

where D^E

$$D^E(x) = \frac{1}{(2\pi)^3} \int e^{i(kx - |kx - x_0| + t)} \frac{d^3k}{\omega}. \quad (16.26)$$

In addition to the chronological product, we shall also define the so-called ordered (or normal) product of $A_\mu(x)$ and $A_\nu(x')$. In this product, denoted by $N(A_\mu(x) A_\nu(x'))$, the photon emission operators $c_{\lambda\lambda'}$ are on the left and the absorption operators c_{λ} are on the right:

$$N(A_\mu(x) A_\nu(x')) = \frac{1}{2V} \sum_{k_1, k'_1} \frac{1}{\sqrt{\omega\omega'}} \{ c_k c_{k'} e^{i(kx+k'x')} + \quad (16.27)$$

¹⁾ R. Feynman, Phys. Rev. 76, 769 (1949).

We note that the vacuum expectation value of $N(A_\mu(x) A_\nu(x'))$ vanishes.

Comparing (16.27) and (16.23) for the ordered and chronological products of $A_\mu(x)$ and $A_\nu(x')$, and using the quantum conditions (15.23), we arrive at

$$P(A_\mu(x) A_\nu(x')) - N(A_\mu(x) A_\nu(x')) = \frac{1}{2} D^\nu(x - x') \delta_{\mu\nu}. \quad (16.28)$$

This relation will be of importance in investigating the scattering matrix (see Section 22).

Let us examine the function $\underline{D}^E(x)$ in more detail. Integrating (16.26) over the angle between $\underline{k}$ and $\underline{x}$, we can write $\underline{D}^E$ in the form

$$D^\nu(x) = \frac{1}{4\pi i} [\delta_+(x - |t|) - \delta_+(-x - |t|)], \quad (16.26')$$

where

$$\delta_+(x) \equiv \frac{1}{\pi} \int_0^\infty e^{ixt} dt = \delta(x) + \frac{i}{\pi x}. \quad (16.29)$$

The last equation means that integrals containing δ_+ in the form $\int_{-\infty}^\infty f(x) \delta_+(x) dx$ (where $f(x)$ is a function with no poles on the real axis and which behaves at infinity so that the integral exists) are to be calculated according to

$$\int_{-\infty}^\infty f(x) \delta_+(x) dx = \int_0^\infty f(x) \frac{dx}{x} = f(0) + \frac{i}{\pi} P \int_{-\infty}^\infty f(x) \frac{dx}{x}, \quad (16.30)$$

where the contour of integration C passes above the pole at $x = 0$, and P is the Cauchy principal part of the integral. This follows immediately from the definition of the function δ_+ . In fact,

$$\begin{aligned} \int_{-\infty}^\infty f(x) \delta_+(x) dx &= \frac{1}{\pi} \lim_{A \rightarrow \infty} \int_{-A}^A dx \left\{ f(x) \int_0^A e^{ixt} dt \right\} = \\ &= \frac{1}{\pi} \lim_{A \rightarrow \infty} \int_{-A}^A f(x) \frac{e^{iAx} - 1}{ix} dx. \end{aligned}$$

If we displace the contour of integration into the upper half-plane and go to the limit $A \rightarrow \infty$, the integral with the term e^{iAx} vanishes, and we obtain the first part of (16.30). Further displacing the contour of integration from the real axis, and passing above the pole $x = 0$ in the remaining integral, we obtain the second part of (16.30).

Using (16.29), $\underline{D}^E$ may be written

$$\begin{aligned} D^\nu(x) &= \frac{1}{(2\pi)^4 i} \left[\pi \delta(r - |t|) + \frac{i}{r - |t|} - \pi \delta(r + |t|) + \frac{i}{r + |t|} \right] = \\ &= \frac{2}{(2\pi)^4 i} \left[\pi \delta(r^2 - t^2) + \frac{i}{r^2 - t^2} \right] = \frac{1}{2\pi i} \delta_+(x^2). \end{aligned} \quad (16.31)$$

From this it can be seen that $\underline{D}^E$ is invariant under Lorentz transformation, and satisfies the wave equation when $x^2 \neq 0$.

We shall later need to Fourier analyze $\underline{D}^E$, and therefore, let us consider

$$\begin{aligned} D^\nu(x) &= \int D^\nu(p) e^{ipx} d^4p, \\ D^\nu(p) &= \frac{1}{(2\pi)^4} \int D^\nu(x) e^{-ipx} d^4x. \end{aligned} \quad (16.31')$$

Inserting (16.26) into this and noting that

$$\frac{1}{(2\pi)^4} \int e^{i(k-p)x} d^4x = \delta(k - p),$$

we obtain

$$D^\nu(p) = \frac{1}{(2\pi)^4} \int \frac{d^4k}{\omega} \left\{ \delta(k - p) \int_{-\infty}^\infty e^{-it(\omega + |k - p|)} dt \right\}.$$

The integral in the braces can, according to (16.29), be written

$$\begin{aligned} \int_{-\infty}^\infty e^{it(p_0 - \omega + |k - p|)} dt &= \int_0^\infty e^{it(p_0 - \omega)} dt + \int_0^\infty e^{-it(p_0 + \omega)} dt = \\ &= \pi [\delta_+(p_0 - \omega) + \delta_+(-p_0 - \omega)]. \end{aligned}$$

Therefore,

$$\begin{aligned} \langle p \rangle &= \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \delta_+(p_0 - |p|) + \delta_+(-p_0 - |p|) = \\ &= \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \left[\pi \delta(p_0 - |p|) + \frac{i}{p_0 - |p|} + \pi \delta(p_0 + |p|) - \frac{i}{p_0 + |p|} \right] = \\ &= \frac{2}{(2\pi)^4} \left[\pi \delta(p^2 - p_0^2) - \frac{i}{p^2 - p_0^2} \right] = \frac{1}{(2\pi)^4} \delta_+(p^2 - p_0^2), \end{aligned} \quad (16.32)$$

where $p^2 = p^2 - p_0^2$.

In integrating over p_0 in (16.31') we must make use of (16.30). This can be done either by integrating along the real p_0 axis and using the principal value of the integral, or by dropping the δ -function in (16.29) and integrating along a contour passing above the pole $p_0 = |p|$ and below the pole $p_0 = -|p|$. This procedure can be formulated differently; instead of passing above the pole $p_0 = |p|$, we may move this pole into the lower half-plane, and similarly, instead of passing below the pole $p_0 = -|p|$, we may move this pole into the upper half-plane. This procedure can be formalized in the following way: in integrating (16.31') over p_0 , the space part p^2 should be replaced by $p^2 - i\epsilon$, where ϵ is an infinitesimal positive number, and the contour of integration should be closed in the lower half-plane if $t > 0$, and in the upper half-plane if $t < 0$ (see Fig. 4).

With this rule we can write $\underline{D}^E$ in the form

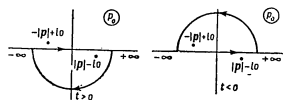


Fig. 4

$$D^E(p) = \frac{2}{i} \frac{1}{(2\pi)^4} \frac{1}{p^2 - i\epsilon}, \quad p^2 \rightarrow p^2 - i\epsilon. \quad (16.33)$$

We note that $\underline{D}^E(x)$ is continuous for $t = 0$ and $t \neq 0$, whereas the derivative $\frac{\partial}{\partial t} \underline{D}^E(x)$ at $t = 0$ has a discontinuity of $-\frac{2i}{V} \sum e^{i p x} = -2i\delta(x)$. Therefore, at $t = 0$,

$$\frac{\partial^2}{\partial t^2} D^E(x) = -2i\delta(x) \delta(t) = -2i\delta(x).$$

For $t \neq 0$, we have according to (16.28)

$$\square D^E(x) = 0,$$

and, therefore, for all values of t , $\underline{D}^E$ satisfies the equation

$$\square D^E(x) = 2i\delta(x). \quad (16.34)$$

It is easy to go to (16.33) from this equation, bearing in mind that

$$\delta(x) = \frac{1}{(2\pi)^4} \int e^{i p x} d^4 p.$$

We note that from (16.4), (16.22), and (16.26) it follows that

$$D^E(x) = D^{(1)}(x) - i\epsilon D(x), \quad (16.35)$$

where $\epsilon = \pm 1$ for $t \lesseqgtr 0$.

§ 17. Quantization of the Electron-Positron Field.

1. Variational Principle for the Dirac Equation. Energy-Momentum Tensor.

In Chapter II we studied the properties of the individual electron. To study an arbitrary system of non-interacting electrons and positrons we may, just as in the case of a system of photons, go over to an occupation-number representation. For a system of photons, as was shown in Section 15, the transition to this representation is equivalent to quantizing the electromagnetic field. For a system of electrons, the transition to the occupation-number representation is equivalent to quantizing the electron-positron field, that is, the field given by the Dirac wave functions ψ and $\bar{\psi}$. In doing this, the Dirac equation is considered a field equation like Maxwell's equations, rather than the equation of a single particle; quantization of the electron-positron field means that the wave functions ψ and $\bar{\psi}$ are considered operators which operate in occupation-number space and satisfy definite commutation rules. The dependence of these operators on space and time is given by the Dirac equation.

We note that the Dirac equation can be obtained from a variational principle

$$\delta \int L d^4 x = 0,$$

if the Lagrangian is taken as

$$L = -\frac{1}{2} \left(\bar{\psi} \gamma_\mu \frac{\partial \psi}{\partial x_\mu} - \frac{\partial \bar{\psi}}{\partial x_\mu} \gamma_\mu \psi \right) - m \bar{\psi} \psi. \quad (17.1)$$

In the variation of L , the functions ψ and $\bar{\psi}$ are to be considered independent.

Having an expression for the Lagrangian, we can use the general equations of field theory (see Section 49) to define the energy-momentum tensor and the current vector of the electron-positron field. The energy-momentum tensor $T_{\mu\nu}$ is given by

$$T_{\mu\nu} = - \sum_r \frac{\partial L}{\partial \dot{\psi}_r} \frac{\partial \dot{\psi}_r}{\partial x_\nu} + L \delta_{\mu\nu} = \frac{1}{2} \left(\bar{\psi} \gamma_\mu \frac{\partial \psi}{\partial x_\nu} - \frac{\partial \bar{\psi}}{\partial x_\mu} \gamma_\nu \psi \right) \quad (17.2)$$

and satisfies the continuity equation

$$\frac{\partial}{\partial x_\mu} T_{\mu\nu} = 0.$$

The energy density and momentum density are given, respectively, by

$$\left. \begin{aligned} w &= -T_{44} = \frac{1}{2i} \left(-\psi^* \frac{\partial \psi}{\partial t} + \frac{\partial \psi^*}{\partial t} \psi \right), \\ \mathbf{g} &= \frac{1}{2i} \left(\psi^* \frac{\partial \psi}{\partial x} - \frac{\partial \psi^*}{\partial x} \psi \right). \end{aligned} \right\} \quad (17.3)$$

The total energy and total momentum of the electron-positron field are given by the integrals of w and $\mathbf{g}$ over all three-space, namely

$$\left. \begin{aligned} E &= \int w d^3x = -\frac{1}{i} \int \bar{\psi} \gamma_4 \frac{\partial \psi}{\partial t} d^3x, \\ P &= \int \mathbf{g} d^3x = \frac{1}{i} \int \bar{\psi} \gamma_4 \frac{\partial \psi}{\partial x} d^3x. \end{aligned} \right\} \quad (17.3')$$

The tensor $T_{\mu\nu}$ is not symmetric. It is, however, possible to form a symmetric tensor

$$\theta_{\mu\nu} = \frac{1}{2} (T_{\mu\nu} + T_{\nu\mu}),$$

which satisfies the same continuity equation

$$\frac{\partial \theta_{\mu\nu}}{\partial x_\nu} = 0$$

and leads to the same total energy and momentum as does $T_{\mu\nu}$.

Let us now find the current density of the electron-positron field. It is defined by the general equation (see Section 49)

$$j_\mu = -ie \sum_r \left(\frac{\partial L}{\partial \dot{\psi}_r} \dot{\psi}_r - \dot{\psi}_r^* \frac{\partial L}{\partial \dot{\psi}_r^*} \right).$$

Using the Lagrangian of (17.1), we must consider the variables ψ and $\bar{\psi}$, not ψ and ψ^* , as independent; therefore, the current four-vector is obtained in the form¹⁾

$$j_\mu = -ie \left(\frac{\partial L}{\partial \dot{\psi}} \dot{\psi} - \bar{\psi} \frac{\partial L}{\partial \dot{\bar{\psi}}} \right) = ie \bar{\psi} \gamma_\mu \psi. \quad (17.4)$$

(The last term is an abbreviated form of writing $e i \bar{\psi} \alpha (\gamma_\mu) \alpha \beta \psi$.) The charge density and the current three-vector are given by

$$\rho = \frac{1}{i} j_4 = e \bar{\psi} \psi, \quad \mathbf{j} = ie \bar{\psi} \boldsymbol{\gamma} \psi = e \bar{\psi} \boldsymbol{\alpha} \psi. \quad (17.5)$$

The total charge of the field is

$$Q = e \int \bar{\psi} \gamma_4 \psi d^3x. \quad (17.5')$$

We see that the charge density is positive definite (compare Section 52).

2. Quantum Conditions for the Electron-Positron Field.

Let us now go on to a study of the quantum conditions of the electron-positron field. As has been mentioned above, we must consider the field components ψ operators acting in occupation-number space and satisfying certain definite quantum conditions. These quantum conditions, however, differ from those for the electromagnetic field as given by (15.23). Indeed, it follows from (15.23) that the number of particles in any given state may be arbitrary, whereas the Dirac equation describes spin $1/2$ particles, which behave according to Fermi-Dirac statistics, so that the number of particles in a given state may be either one or zero (the Pauli exclusion principle).

In order to quantize the electron-positron field we shall expand the general solution of the Dirac equation (7.30) in the orthonormal set of functions $\psi_n^{(+)}$ and $\psi_n^{(-)}$, where $\psi_n^{(+)}$ and $\psi_n^{(-)}$ are the solutions with positive and negative frequencies, respectively:

¹⁾ This definition of the current differs from that in Section 7 by the factor e .

$$\left. \begin{aligned} \psi &= \sum_n (a_n \psi_n^{(+)} + b_n^+ \psi_n^{(-)}) \\ \bar{\psi} &= \sum_n (a_n^+ \bar{\psi}_n^{(+)} + b_n \bar{\psi}_n^{(-)}) \end{aligned} \right\} \quad (17.6)$$

where the a_n and a_n^+ , as well as the b_n and b_n^+ are considered Hermitian conjugate operators satisfying the following quantum conditions:

$$\left. \begin{aligned} \{a_r, a_s^+\} &= \delta_{rs}, & \{b_r, b_s^+\} &= \delta_{rs}, \\ \{a_r, a_s\} &= 0, & \{a_r^+, a_s^+\} &= 0, \\ \{b_r, b_s\} &= 0, & \{b_r^+, b_s^+\} &= 0, \\ \{a_r, b_s\} &= 0, & \{a_r, b_s^+\} &= 0, \\ \{a_r^+, b_s\} &= 0, & \{a_r^+, b_s^+\} &= 0, \end{aligned} \right\} \quad (17.7)$$

where $\{AB\} = AB + BA$ (the bracket $\{AB\}$ is called the anticommutator bracket of the operators A and B).

When no external fields are present, we can expand ψ and $\bar{\psi}$ in plane waves which are eigenstates of the momentum and polarization. We shall write this expansion in the form

$$\left. \begin{aligned} \psi_\alpha(x) &= \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} \sum_{\mathbf{r}} (a_{\mathbf{r}}(\mathbf{p}) u_{\alpha}^{\mathbf{r}}(\mathbf{p}) e^{i\mathbf{p} \cdot \mathbf{x}} + b_{\mathbf{r}}^+(\mathbf{p}) \bar{u}_{\alpha}^{\mathbf{r}}(\mathbf{p}) e^{-i\mathbf{p} \cdot \mathbf{x}}), \\ \bar{\psi}_\alpha(x) &= \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} \sum_{\mathbf{r}} (a_{\mathbf{r}}^+(\mathbf{p}) \bar{u}_{\alpha}^{\mathbf{r}}(\mathbf{p}) e^{-i\mathbf{p} \cdot \mathbf{x}} + b_{\mathbf{r}}(\mathbf{p}) u_{\alpha}^{\mathbf{r}}(\mathbf{p}) e^{i\mathbf{p} \cdot \mathbf{x}}), \end{aligned} \right\} \quad (17.8)$$

where V is the normalizing volume and summation over $\mathbf{r}$ (for the values $\mathbf{r} = 1, 2$) indicates summation over the two spin states; $u_{\alpha}^{\mathbf{r}}$ and $\bar{u}_{\alpha}^{\mathbf{r}}$ ($\mathbf{r} = \mathbf{r}^{\mathbf{r}}, \mathbf{r}^{\mathbf{r}}$) are constant spinors satisfying the following orthogonality and normalization conditions:

$$\sum_{\mathbf{r}} u_{\alpha}^{\mathbf{r}}(\mathbf{p})^* u_{\beta}^{\mathbf{r}}(\mathbf{p}) = \delta_{\alpha\beta}, \quad \sum_{\mathbf{r}} \bar{u}_{\alpha}^{\mathbf{r}}(\mathbf{p})^* \bar{u}_{\beta}^{\mathbf{r}}(\mathbf{p}) = \delta_{\alpha\beta}. \quad (17.9)$$

In the scalar product $\mathbf{p} \cdot \mathbf{x} = \mathbf{p} \cdot \mathbf{x} - p_4 x_4$, the fourth component of $\mathbf{p}$ is defined as $p_4 = i p_0 \equiv i E = i \sqrt{\mathbf{p}^2 + m^2}$. The spinors $\bar{u}^{\mathbf{r}}$ and $\bar{v}^{\mathbf{r}}$ are related to $u^{\mathbf{r}}$ and $v^{\mathbf{r}}$ by

$$\bar{u}^{\mathbf{r}} = u^{\mathbf{r}*} \beta, \quad \bar{v}^{\mathbf{r}} = v^{\mathbf{r}*} \beta.$$

Without loss of generality we may consider the solution $\bar{\psi} = \bar{\psi} e^{-i\mathbf{p} \cdot \mathbf{x}}$ to be the charge conjugate of $\psi e^{i\mathbf{p} \cdot \mathbf{x}}$, so that [see (8.30)]

$$\bar{\psi} = \bar{u} C', \quad \bar{\psi} = C'^{-1} u \quad (17.10)$$

(this is true because C is a unitary matrix and, therefore, the orthonormality conditions for $\bar{\psi}$ follow from those for ψ).

We note also the easily obtained relation

$$\sum_{\mathbf{r}} u_{\mathbf{r}}^* u_{\mathbf{s}}^{\mathbf{r}} = \delta_{\mathbf{rs}}, \quad (\mathbf{r}, \mathbf{s} = 1, 2, 3, 4), \quad (17.11)$$

where $\bar{\mathbf{r}} = \mathbf{r}^{\mathbf{r}}, \bar{\mathbf{r}} = \mathbf{r}^{\mathbf{r}}$.

From the quantum conditions (17.7) it follows that the eigenvalues of

$$N_{\mathbf{r}}^+ = a_{\mathbf{r}}^+ a_{\mathbf{r}}, \quad N_{\mathbf{r}}^- = b_{\mathbf{r}}^+ b_{\mathbf{r}} \quad (17.12)$$

are either zero or one, and that the nonvanishing matrix elements of the operators $a_{\mathbf{r}}, a_{\mathbf{r}}^+, b_{\mathbf{r}}, b_{\mathbf{r}}^+$ in the representation in which $N_{\mathbf{r}}^+$ and $N_{\mathbf{r}}^-$ are diagonal are given by¹⁾

$$\left. \begin{aligned} (a_{\mathbf{r}})_{N_{\mathbf{r}}^+ - 1, N_{\mathbf{r}}^+} &= \sqrt{N_{\mathbf{r}}^+}, & (b_{\mathbf{r}})_{N_{\mathbf{r}}^- - 1, N_{\mathbf{r}}^-} &= \sqrt{N_{\mathbf{r}}^-}, \\ (a_{\mathbf{r}}^+)_{N_{\mathbf{r}}^+, N_{\mathbf{r}}^+ - 1} &= \sqrt{N_{\mathbf{r}}^+}, & (b_{\mathbf{r}}^+)_{N_{\mathbf{r}}^-, N_{\mathbf{r}}^- - 1} &= \sqrt{N_{\mathbf{r}}^-}, \end{aligned} \right\} \quad (17.13)$$

and we may, therefore, say that the operators $a_{\mathbf{r}}$ and $b_{\mathbf{r}}$ decrease, whereas the $a_{\mathbf{r}}^+$ and $b_{\mathbf{r}}^+$ increase, the numbers $N_{\mathbf{r}}^+$ and $N_{\mathbf{r}}^-$ by unity.

We shall now show that the quantum conditions (17.7) lead to the correct corpuscular picture. For this purpose let us determine the energy, momentum, and charge of the field. Like the field components, these quantities are operators. We shall determine them with the aid of (17.3'), (17.5'), and the expansion (17.8), where the $a_{\mathbf{r}}, a_{\mathbf{r}}^+, b_{\mathbf{r}}, b_{\mathbf{r}}^+$ satisfy the quantum conditions (17.7).

Using the normalization conditions (17.9), which are valid for all $\mathbf{p}$, it is easy to show that

¹⁾ This is easily shown by considering $\begin{pmatrix} a_{\mathbf{r}}^+ a_{\mathbf{r}} & a_{\mathbf{r}}^+ a_{\mathbf{r}}^+ \\ a_{\mathbf{r}} a_{\mathbf{r}} & a_{\mathbf{r}} a_{\mathbf{r}}^+ \end{pmatrix}$ and using (17.7) together with the diagonality of $a_{\mathbf{r}}^+ a_{\mathbf{r}}$.

$$\left. \begin{aligned} \mathcal{E} &= \sum_p \sum_{r=1}^2 \epsilon (a_r^\dagger a_r - b_r b_r^\dagger), \\ P &= \sum_p \sum_{r=1}^2 p (a_r^\dagger a_r - b_r b_r^\dagger), \\ Q &= e \sum_p \sum_{r=1}^2 (a_r^\dagger a_r + b_r b_r^\dagger). \end{aligned} \right\} \quad (17.14)$$

Inserting (17.12) and making use of (17.7), we obtain

$$\left. \begin{aligned} \mathcal{E} &= \sum_p \sum_{r=1}^2 \epsilon (N_r^+ + N_r^- - 1), \\ P &= \sum_p \sum_{r=1}^2 p (N_r^+ + N_r^- - 1), \\ Q &= e \sum_p \sum_{r=1}^2 (N_r^+ - N_r^- + 1). \end{aligned} \right\} \quad (17.14')$$

Expressions (17.14) and (17.14') for the energy are not positive definite. In other words, the energy of the electron-positron field, both in the classical theory and when the a and b satisfy the quantum conditions (17.7), can have both positive and negative values.

We shall now show that the vacuum state of the electron-positron field can be defined so that the energy of the field is positive definite.

Let us define the vacuum state as that in which the energy has the lowest possible value, that is as the state in which all the $N_{\underline{r}}^+$ and $N_{\underline{r}}^-$ are zero, or in other words¹⁾ when

$$\left. \begin{aligned} \langle a_r^\dagger a_r \rangle_0 &= 0, & \langle a_r a_r^\dagger \rangle_0 &= 1, \\ \langle b_r^\dagger b_r \rangle_0 &= 0, & \langle b_r b_r^\dagger \rangle_0 &= 1. \end{aligned} \right\} \quad (17.15)$$

According to (17.14') the energy and the charge in the vacuum state are given by

¹⁾ The symbol $\langle L \rangle_0$ denotes the expectation value of L in the vacuum state.

$$\left. \begin{aligned} \mathcal{E}_0 &= - \sum_p \sum_{r=1}^2 \epsilon, \\ P_0 &= \sum_p \sum_{r=1}^2 p, \\ Q_0 &= \sum_p \sum_{r=1}^2 e. \end{aligned} \right\} \quad (17.16)$$

Therefore, the energy, momentum, and charge of the quantized electron field differ from their values in the vacuum by

$$\left. \begin{aligned} \mathcal{E} &= \sum_p \sum_{r=1}^2 \epsilon (N_r^+ + N_r^-), \\ P &= \sum_p \sum_{r=1}^2 p (N_r^+ + N_r^-), \\ Q &= \sum_p \sum_{r=1}^2 e (N_r^+ - N_r^-). \end{aligned} \right\} \quad (17.17)$$

We see that the numbers $N_{\underline{r}}^+$ and $N_{\underline{r}}^-$ which can take on the values 0 and 1 enter as sums in the expressions for the energy and momentum and as differences in the expression for the total charge. These numbers are, therefore, interpreted as the numbers of electrons and positrons having energy ϵ , momentum $\underline{p}$, and a definite spin orientation.

Thus, the quantum conditions (17.7) lead indeed to the correct corpuscular picture, since the energy, momentum, and the charge of the field is given by the sum of the energies, momenta, and charges of the separate electrons and positrons.

If the Dirac equation (7.30) is considered a wave equation determining the various states of an individual electron, Equation (17.16) can be interpreted in the following way: the vacuum is the state in which all the negative energy levels are filled. This infinite "negative background" of electrons is not in itself observed, but under the influence of various external fields the electron can undergo a transition from a negative energy state to a positive energy state (this necessitates, obviously, an energy no less than $2mc^2$), leaving a "hole" in the infinite negative sea, and this hole behaves like a particle with a positive energy and a charge with the opposite sign but the same magnitude as the electron charge.

Such a "hole" in the infinite sea of negative energy electrons can be interpreted as a positron, and the creation of a "hole" represented by an electron going from a negative to a positive energy state can be interpreted in terms of electron-positron pair creation.¹⁾

In quantizing the electron-positron field we started with the expansion (17.6) of a general solution of the Dirac equation in a series of the orthonormal set of functions $\psi_{\underline{p}}^{(+)}$ and $\psi_{\underline{p}}^{(-)}$ corresponding to positive and negative frequencies. This expansion is possible in the absence of external fields (we have considered this in detail already) and when the external fields are sufficiently weak.

¹⁾ It is, of course, possible to treat electrons as "holes" in the infinite positron sea. Later we shall go through a more detailed investigation of the symmetry of the theory with respect to replacing $\underline{e}$ by $-\underline{e}$.

In a very strong external field we arrive at the difficulty of separating the frequency spectrum into "positive" and "negative" frequencies, as was pointed out at the end of Section 11. If the external field is varied adiabatically, for instance, by varying the potential well depth V_0 in the example of Section 11, the frequency ω_k which for $V_0 = 0$ belongs to the "upper continuum" ($\omega > \bar{m}$), moves into the "lower continuum" ($\omega < -\bar{m}$) for some value $V_0 > V_0^{(k)}$. When the frequency ω_k crosses the boundary of the "lower continuum" we must start considering the state corresponding to this frequency a positron state. Therefore, in the series expansion for the field operator ψ given by (17.6), the operator a_n corresponding to this frequency (we shall denote it by a_k) becomes an emission operator instead of an absorption operator. We shall denote this operator by b_k^+ , so that

$$a_k \rightarrow b_k^+, \quad V_0 > V_0^{(k)}.$$

When $V_0 < V_0^{(k)}$, while each electron and positron state can be associated adiabatically with some free electron or free positron state, the wave function Φ_0 in the vacuum state satisfies

$$a_n \Phi_0 = 0, \quad b_n \Phi_0 = 0. \quad (17.15')$$

This means according to (17.15) that in the vacuum state all the occupation numbers vanish:

$$\Phi_0 = (0, 0, \dots, 0, \dots; 0, 0, \dots, 0, \dots), \quad V_0 < V_0^{(k)}.$$

When $V_0 > V_0^{(k)}$, (17.15') are replaced by

$$\left. \begin{aligned} a_n \Phi_0 &= 0, \quad n \neq k, \\ b_n \Phi_0 &= 0, \\ b_k^+ \Phi_0 &= 0 \end{aligned} \right\} \quad V_0 > V_0^{(k)} \quad (17.15'')$$

($n = 0$ corresponds to the frequency ω_k). These conditions mean that the vacuum now represents that state in which one of the positron levels is filled, namely

$$\Phi_0 = (0, 0, \dots, 0, \dots; 0, 0, \dots, 1^-, 0, \dots), \quad V_0 > V_0^{(k)}.$$

If there is an electron in the field, then the wave function for $V_0 < V_0^{(k)}$ is of the form

$$\Phi_1 = (0, 0, \dots, 1^+, \dots; 0, 0, \dots), \quad V_0 < V_0^{(k)},$$

and for $V_0 > V_0^{(k)}$ it is $\Phi_1 = (0, 0, \dots, 1^+, \dots; 0, 0, \dots, 1^-, \dots)$, $V_0 > V_0^{(k)}$.

Thus, the behavior of an electron in a strong field $V_0 > V_0^{(k)}$ is equivalent to the behavior of an electron in the presence of a positron.

As will be shown later, electrons and positrons can decay into photons or be absorbed by an external field.¹⁾ We, therefore, see that an electron in a strong field ($V_0 > V_0^{(k)}$) can be absorbed by the field.

This result is simple to interpret in the language of hole theory. We have said above that the vacuum state is defined as that state in which all the levels with $\omega < -\bar{m}$ are filled. The occurrence of a new unfilled level with $\omega_k < -\bar{m}$ in the field $V_0 > V_0^{(k)}$ means that an electron can go over to this lower level in this field, so that it is absorbed.

§ 18. Anticommutators of the Electron-Positron Field. The Singular Functions

$$\Delta(x), \Delta^{(1)}(x), \Delta^F(x), \Delta^R(x).$$

1. Quantum Conditions for the Operators $\psi, \bar{\psi}$

Let us find the quantum conditions satisfied by the field operators $\psi, \bar{\psi}$. For this purpose let us use the quantum conditions (17.7) satisfied by the a, a^+, b, b^+ .

It follows from (17.7) that

$$\{\psi(x), \psi(x')\} = 0, \quad \{\bar{\psi}(x), \bar{\psi}(x')\} = 0. \quad (18.1)$$

In addition,

$$\{\psi(x), \bar{\psi}(x')\} = \frac{1}{V} \sum_p \sum_{r=1}^2 u_p^r \bar{u}_p^r e^{-ip(x-x')} + \frac{1}{V} \sum_p \sum_{r=1,2} v_p^r \bar{v}_p^r e^{-ip(x-x')}, \quad (18.2)$$

where $\underline{p} = \underline{p} = p_0, \underline{p}$. Let us determine the functions

$$\left. \begin{aligned} \frac{1}{V} \sum_p \sum_{r=1}^2 u_p^r \bar{u}_p^r e^{-ip(x-x')} &= -iS_{\psi\bar{\psi}}^+(x-x'), \\ \frac{1}{V} \sum_p \sum_{r=1}^2 v_p^r \bar{v}_p^r e^{-ip(x-x')} &= -iS_{\psi\bar{\psi}}^-(x-x'), \end{aligned} \right\} \quad (18.3)$$

We shall first find $\sum_{r=1}^2 u_p^r \bar{u}_p^r$ and $\sum_{r=1}^2 v_p^r \bar{v}_p^r$. We introduce the operator

¹⁾ See Sections 33 and 40.

$$\eta_+ = \frac{1}{2\epsilon}(\alpha p + m\beta + \epsilon), \quad \epsilon = +\sqrt{p^2 + m^2}. \quad (18.4)$$

Since the spinors $\psi^1, \psi^2, \psi^3, \psi^4$ belong to states with negative frequencies, then obviously

$$\eta_+ u^r = \begin{cases} u^r, & r = 1, 2, \\ 0, & r = 3, 4, \end{cases} \quad (18.5)$$

and, therefore,

$$\sum_{r=1}^2 u_s^r u_t^{r*} = \sum_{r=1}^4 \frac{1}{2\epsilon}(\alpha p + m\beta + \epsilon)_{st} u_s^r u_t^{r*}.$$

But according to (17.11), $\sum_{r=1}^4 u_s^r u_t^{r*} = \delta_{st}$ and, therefore,

$$\sum_{r=1}^2 u_s^r u_t^{r*} = \frac{1}{2\epsilon}(\alpha p + m\beta + \epsilon)_{st}.$$

Multiplying this equation on the right by $(\gamma_\lambda)_{\rho\lambda}$ and bearing in mind the relation between the α and γ matrices, we finally obtain

$$\sum_{r=1}^2 u_s^r \bar{u}_\lambda^r = -\frac{1}{2\epsilon}(i\gamma_\lambda p_\lambda - m)_{s\lambda}. \quad (18.6)$$

The sum $\sum_{r=1}^2 \bar{u}_\lambda^r \psi_\lambda^r$ can be found in a similar way. We introduce the operator

$$\eta_- = -\frac{1}{2\epsilon}(-\alpha p + m\beta - \epsilon), \quad \epsilon = +\sqrt{p^2 + m^2}, \quad (18.7)$$

¹⁾ Actually these are bispinors; see Section 7.

and in the same way as previously we obtain¹⁾

$$\sum_{r=1}^2 \bar{u}_\lambda^r \psi_\lambda^r = \frac{1}{2\epsilon}(-i\gamma_\lambda p_\lambda - m)_{\lambda\lambda}. \quad (18.8)$$

Let us now find $\bar{S}_{\alpha\beta}^+$ and $\bar{S}_{\alpha\beta}^-$. Using (18.6) and (18.8) we obtain

$$\begin{aligned} -iS_{\alpha\beta}^+(x) &= -\frac{1}{V} \sum_p \frac{1}{2\epsilon} (i\gamma_\alpha p_\alpha - m)_{\alpha\beta} e^{ipx} = -\left(\gamma_\alpha \frac{\partial}{\partial x_\alpha} - m\right)_{\alpha\beta} \frac{1}{V} \sum_p \frac{1}{2\epsilon} e^{ipx}, \\ -iS_{\alpha\beta}^-(x) &= \frac{1}{V} \sum_p \frac{1}{2\epsilon} (-i\gamma_\alpha p_\alpha - m)_{\alpha\beta} e^{-ipx} = \left(\gamma_\alpha \frac{\partial}{\partial x_\alpha} - m\right)_{\alpha\beta} \frac{1}{V} \sum_p \frac{1}{2\epsilon} e^{-ipx}. \end{aligned} \quad (18.9)$$

Introducing the two functions²⁾

²⁾ Equations (18.6) and (18.8) can be obtained also from the results of Section 9, paragraph 3. We write

$$\begin{aligned} \sum_{r=1}^2 u_s^r \bar{u}_\lambda^r &= \sum_{r=1}^2 u_s^{r,A} \bar{u}_\lambda^{r,A} = \sum_{r=1}^2 \left(\frac{m - i\gamma_5 p_5}{2\Delta} u^{r,A} \right)_s \bar{u}_\lambda^{r,A} = \\ &= \sum_{r=1}^2 \left(\frac{m - i\gamma_5 p_5}{2\Delta} \right)_{sr} \bar{u}_\lambda^{r,A} u_s^{r,A}, \end{aligned}$$

and since according to (19.19) (with the substitutions $\mu \rightarrow r, \nu \rightarrow \gamma, \psi \rightarrow \bar{\psi}, u_{\mu\lambda}(\psi) \rightarrow u_{\gamma}^{r,A}, \bar{u}_{\mu\lambda}(\psi) \rightarrow \bar{u}_{\gamma}^{r,A}$)

$$\sum_{r=1}^2 \bar{u}_\lambda^{r,A} u_s^{r,A} = \delta_{\lambda\beta},$$

we obtain

$$\sum_{r=1}^2 u_s^r \bar{u}_\lambda^r = -\frac{1}{2\epsilon}(i\gamma_\lambda p_\lambda - m)_{s\lambda}.$$

Equation (18.8) is obtained similarly.

²⁾ P. Dirac, Proc. Cambridge Phil. Soc. 30, 150 (1934).

$$\left. \begin{aligned} \Delta(x) &= \frac{1}{V} \sum_p e^{ipx} \frac{\sin \epsilon t}{\epsilon} = \frac{1}{(2\pi)^3} \int e^{ipx} \frac{\sin \epsilon t}{\epsilon} d^3p, \\ \Delta^{(1)}(x) &= \frac{1}{V} \sum_p e^{ipx} \frac{\cos \epsilon t}{\epsilon} = \frac{1}{(2\pi)^3} \int e^{ipx} \frac{\cos \epsilon t}{\epsilon} d^3p, \end{aligned} \right\} \quad (18.10)$$

we write $-iS_{\alpha\beta}^+$ and $-iS_{\alpha\beta}^-$ in the form

$$\left. \begin{aligned} -iS_{\alpha\beta}^+(x) &= -\frac{1}{2} \left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m \right)_{\alpha\beta} [\Delta^{(1)}(x) - i\Delta(x)], \\ -iS_{\alpha\beta}^-(x) &= \frac{1}{2} \left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m \right)_{\alpha\beta} [\Delta^{(1)}(x) + i\Delta(x)]. \end{aligned} \right\} \quad (18.11)$$

Using these formulas, we obtain the following expression for the anticommutator $\{\psi_\alpha(x), \bar{\psi}_\beta(x')\}$:

$$\{\psi_\alpha(x), \bar{\psi}_\beta(x')\} = -iS_{\alpha\beta}(x - x'), \quad (18.12)$$

where

$$S_{\alpha\beta}(x) = S_{\alpha\beta}^+(x) + S_{\alpha\beta}^-(x) = -\left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m \right)_{\alpha\beta} \Delta(x). \quad (18.13)$$

The functions $-iS_{\alpha\beta}^+(\underline{x} - \underline{x}')$ and $-iS_{\alpha\beta}^-(\underline{x} - \underline{x}')$ have simple physical meaning; they are the vacuum state expectation values of $\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')$ and $\bar{\psi}_\beta(x') \psi_\alpha(x)$. Indeed, using the expansion (17.8) and relation (17.15) it is easy to show that

$$\left. \begin{aligned} \langle \bar{\psi}_\alpha(x) \bar{\psi}_\beta(x') \rangle_0 &= -iS_{\alpha\beta}^+(x - x'), \\ \langle \bar{\psi}_\beta(x') \psi_\alpha(x) \rangle_0 &= -iS_{\alpha\beta}^-(x - x'). \end{aligned} \right\} \quad (18.14)$$

It follows from this and from (18.11) that

$$\langle \{\psi_\alpha(x), \bar{\psi}_\beta(x')\} \rangle_0 = -S_{\alpha\beta}^{(1)}(x - x'), \quad (18.15)$$

where

$$S_{\alpha\beta}^{(1)}(x) = \left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m \right)_{\alpha\beta} \Delta^{(1)}(x).$$

We note that $\Delta(\underline{x})$ and $\Delta^{(1)}(\underline{x})$ may be considered generalizations of $\underline{D}(\underline{x})$ and $\underline{D}^{(1)}(\underline{x})$ which were introduced in Section 16; if the mass m is set equal to zero in $\Delta(\underline{x})$ and $\Delta^{(1)}(\underline{x})$, we obtain $\underline{D}(\underline{x})$ and $\underline{D}^{(1)}(\underline{x})$.

It is easy to see that $\Delta(\underline{x})$ and $\Delta^{(1)}(\underline{x})$ satisfy the equations

$$(\square - m^2)\Delta(x) = 0, \quad (\square - m^2)\Delta^{(1)}(x) = 0. \quad (18.16)$$

Since $\frac{d^4x}{\epsilon}$ is an invariant, $\Delta(\underline{x})$ and $\Delta^{(1)}(\underline{x})$ are invariant under Lorentz transformations. The function $\Delta(\underline{x})$ vanishes at $t = 0$, and, therefore, also everywhere outside the light cone when $\underline{x}^2 = \underline{x}^2 - t^2 > 0$. The invariance of these functions is also explicitly exhibited when they are written

$$\left. \begin{aligned} \Delta(x) &= \frac{i}{(2\pi)^3} \int e^{ipx} \delta(p^0 + m^2) d^4p, \\ \Delta^{(1)}(x) &= \frac{1}{(2\pi)^3} \int e^{ipx} \delta(p^2 + m^2) d^4p, \end{aligned} \right\} \quad (18.17)$$

where $\epsilon(p) = \frac{p_0}{|p_0|}$.

2. Definition of the Current. Charge Conjugate Operators.

We have seen above that expression (17.4) for the current leads to infinite values of the vacuum charge and current. It is possible to define the current somewhat differently in the quantum theory of the electron-positron field, in particular so that the vacuum charge vanishes. Indeed, according to (18.12) the operators $\bar{\psi}_\alpha(x) \psi_\beta(x')$ and $-\bar{\psi}_\beta(x') \psi_\alpha(x)$ differ only by a c-number, and therefore, the current may be defined as¹⁾

$$\begin{aligned} j_\mu(x) &= \frac{ie}{2} [\bar{\psi}(x), \gamma_\mu \psi(x)] = \frac{ie}{2} (\gamma_\mu)_{\alpha\beta} (\bar{\psi}_\alpha(x) \psi_\beta(x) - \bar{\psi}_\beta(x) \psi_\alpha(x)) = \\ &= \frac{ie}{2} (\bar{\psi}(x) \gamma_\mu \psi(x) - \psi(x) \gamma_\mu \bar{\psi}(x)), \end{aligned} \quad (18.18)$$

where $\tilde{\gamma}_\mu$ is the transpose of γ_μ , that is $(\tilde{\gamma}_\mu)_{\alpha\beta} = (\gamma_\mu)_{\beta\alpha}$. This expression, as well as (17.4), satisfies the continuity equation, but leads to the following expression for the total charge:

1) W. Heisenberg, Z. Physik 90, 209, 692 (1934).

$$Q = -i \int J_4(x) d^3x = \frac{e}{2} \sum_p \sum_{r=1}^2 ([a_r^{\dagger}, a_r] - [b_r^{\dagger}, b_r]) = e \sum_p \sum_{r=1}^2 (N_r^{\dagger} - N_r^-). \quad (18.19)$$

This expression for Q does not contain an infinite vacuum charge, and is identical with the previously obtained expression (17.17).

Let us now introduce the operators

$$\phi' = \bar{\psi} C', \quad \bar{\psi}' = C'^{-1} \psi, \quad (18.20)$$

where the matrix C' is defined by (8.31) and (8.19). We shall call ψ' and $\bar{\psi}'$ the charge conjugates of ψ and $\bar{\psi}$. They satisfy the equations

$$\left. \begin{aligned} (\gamma_4 \frac{\partial}{\partial x_4} + m) \phi' &= 0, \\ \bar{\psi}' (\gamma_4 \frac{\partial}{\partial x_4} - m) &= 0. \end{aligned} \right\} \quad (18.21)$$

Using (8.19) and (8.31) it can be shown that C' is a unitary antisymmetric matrix, that is,

$$C'^* C' = 1, \quad \bar{C}' = -C',$$

and that it satisfies

$$C' (\gamma_4)^* C'^{-1} = -\gamma_4. \quad (18.20')$$

It follows from these expressions that

$$\left. \begin{aligned} \bar{\psi}'_r \bar{\psi} &= \psi (\bar{C}'^{-1})_{rs} \gamma_4 C' \bar{\psi} = \bar{\psi}'_r \psi', \\ \phi \bar{\psi}' &= -\psi (\bar{C}'^{-1})_{rs} C' \bar{\psi} = -\bar{\psi}'_r \phi', \end{aligned} \right\} \quad (18.22)$$

and therefore, (18.18) can be written in a more symmetric form, namely

$$J_4(x) = \frac{ie}{2} (\bar{\psi}'(x) \gamma_4 \psi(x) - \bar{\psi}'(x) \gamma_4 \phi'(x)). \quad (18.23)$$

This expression is invariant under the operation of replacing ψ and $\bar{\psi}$ by their charge conjugates ψ' and $\bar{\psi}'$ and of changing the sign of the charge e .

We shall show that the whole theory of the quantized electron-positron field is invariant with respect to the transformation

$$\psi \rightarrow \psi', \quad \bar{\psi} \rightarrow \bar{\psi}', \quad e \rightarrow -e. \quad (18.24)$$

For this purpose we note that it follows from (17.8) and (17.10) that transformation from $\psi, \bar{\psi}$ to the charge conjugates $\psi', \bar{\psi}'$

$$\left. \begin{aligned} \phi'(x) &= C' \bar{\psi}(x) = \frac{1}{\sqrt{V}} \sum_p \sum_{r=1}^2 [a_r^{\dagger}(p) v^r(p) e^{-ipx} + b_r(p) u^r(p) e^{ipx}], \\ \bar{\psi}'(x) &= C'^{-1} \psi(x) = \frac{1}{\sqrt{V}} \sum_p \sum_{r=1}^2 [a_r(p) \bar{v}^r(p) e^{ipx} + b_r^{\dagger}(p) \bar{u}^r(p) e^{-ipx}] \end{aligned} \right\}$$

is equivalent to the transformation

$$\left. \begin{aligned} a_r &\rightarrow a'_r = b_r, & a_r^{\dagger} &\rightarrow a'^{\dagger}_r = b_r^{\dagger}, \\ b_r &\rightarrow b'_r = a_r, & b_r^{\dagger} &\rightarrow b'^{\dagger}_r = a_r^{\dagger}. \end{aligned} \right\} \quad (18.25)$$

It is easy to see that the operators $a'_r, a'^{\dagger}_r, b'_r, b'^{\dagger}_r$ satisfy the same quantum conditions as do the $a_r, b_r, a_r^{\dagger}, b_r^{\dagger}$, namely

$$\left. \begin{aligned} \{a'_r, a'^{\dagger}_s\} &= \{b'_r, b'^{\dagger}_s\} = \delta_{rs}, \\ \{a'_r, a'_s\} &= \{a'^{\dagger}_r, a'^{\dagger}_s\} = \{b'_r, b'_s\} = \{b'^{\dagger}_r, b'^{\dagger}_s\} = 0. \end{aligned} \right\}$$

Therefore, the anticommutator $\{\psi'_\alpha(x) \bar{\psi}'_\beta(x')\}$ is given by the same equation as the anticommutator $\{\psi_\alpha(x) \bar{\psi}_\beta(x')\}$, namely

$$\{\psi'_\alpha(x) \bar{\psi}'_\beta(x')\} = -iS_{\alpha\beta}(x - x').$$

Furthermore, under (18.25) the vacuum expectation values of $\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')$ do not change, since the primed quantities $a_r', b_r', a_r'^{\dagger}, b_r'^{\dagger}$ satisfy (17.15):

$$\begin{aligned} \langle a_r'^{\dagger} a_r' \rangle_0 &= \langle b_r'^{\dagger} b_r' \rangle_0 = 0, \\ \langle a_r'^{\dagger} a_r'^{\dagger} \rangle_0 &= \langle b_r'^{\dagger} b_r'^{\dagger} \rangle_0 = 1. \end{aligned}$$

Finally, since (18.25) transforms N_r^+ into N_r^- and N_r^- into N_r^+ according to

$$N_r^+ \rightarrow N_r'^+ = N_r^-, \quad N_r^- \rightarrow N_r'^- = N_r^+,$$

it is necessary, in order to leave the current invariant, also to change the sign of $\underline{e}$. Thus, the invariance of the theory under (18.24) has been proved.

3. Chronological Products of the Field Operators.

It will be necessary in the future to have expressions for the various types of products of the electron-positron field components.

First we shall calculate the chronological products (denoted P) introduced in Section 16, in which by definition, the operators are chronologically ordered from left to right. Let us consider the chronological product of $\bar{\psi}_\alpha(x)$ and $\bar{\psi}_\beta(x')$

$$P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) = \begin{cases} \bar{\psi}_\alpha(x) \bar{\psi}_\beta(x'), & t > t', \\ \bar{\psi}_\beta(x') \bar{\psi}_\alpha(x), & t < t'. \end{cases} \quad (18.26)$$

and let us find its vacuum expectation value.

Using (17.8), (17.15), and (18.3), we obtain

$$\begin{aligned} \langle P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) \rangle_0 &= \frac{1}{V} \sum_p \sum_{r=1}^k \bar{u}_{\alpha p} u_{\beta p} e^{ip(x-x')} = -i S_{\beta\alpha}^-(x'-x), \quad t' < t, \\ &= \frac{1}{V} \sum_p \sum_{r=1}^k \bar{u}_{\beta p} u_{\alpha p} e^{ip(x'-x)} = -i S_{\beta\alpha}^+(x'-x), \quad t' > t. \end{aligned} \quad (18.27)$$

Noting that

$$\begin{aligned} -i S_{\beta\alpha}^+(x) &= -\frac{1}{2} \left(\gamma, \frac{\partial}{\partial x} - m \right)_{\beta\alpha} (\Delta^{(0)}(x) - i \Delta(x)), \\ -i S_{\beta\alpha}^-(x) &= \frac{1}{2} \left(\gamma, \frac{\partial}{\partial x} - m \right)_{\beta\alpha} (\Delta^{(0)}(x) + i \Delta(x)), \end{aligned}$$

and making use of (18.10), let us write $\langle P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) \rangle_0$ in the form

$$\langle P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) \rangle_0 = \frac{1}{2} \epsilon(t, t') S_{\beta\alpha}^e(x'-x), \quad (18.28)$$

where¹⁾

$$\begin{aligned} S^e(x) &= \left(\gamma, \frac{\partial}{\partial x} - m \right) \Delta^e(x), \\ \Delta^e(x) &= \frac{1}{(2\pi)^3} \int e^{ipx} \frac{e^{-i\epsilon(t)} d^3p}{\epsilon} = + \sqrt{p^2 + m^2} \end{aligned} \quad (18.29)$$

and $\epsilon(t, t') = +1$ if $t > t'$, and $\epsilon(t, t') = -1$ if $t < t'$.

In a similar way we find that

$$\langle P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) \rangle_0 = \frac{1}{2} \epsilon(t', t) S_{\beta\alpha}^e(x-x'). \quad (18.28')$$

The vacuum expectation values of $P(\bar{\psi}_\alpha \bar{\psi}_\beta)$ and $P(\bar{\psi}_\alpha \bar{\psi}_\beta)$ vanish. We note that $\Delta^E(x)$ may be considered a generalization of the function $\bar{D}^E(x)$ introduced in Section 16; if we set $\underline{m} = 0$ in (18.29), we obtain $\bar{D}^E(x)$.

In calculating the chronological products, we expanded ψ and $\bar{\psi}$ in plane waves. We may, however, start with the general expression (17.16) for ψ in terms of an arbitrary complete orthonormal set of functions. Such an expansion is necessary in investigating the interaction between the electron and the electromagnetic field if the electron is in a stationary external field which cannot be considered a perturbation and must be included in an exact way in the electron wave functions. In this case ψ and $\bar{\psi}$ must be expanded in the eigenfunctions of the electron in the external field under consideration. Going through the same considerations as we used to obtain (18.28), it is easy to show that the vacuum expectation values of $\bar{\psi}$ and ψ are given in general by

$$\langle P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) \rangle_0 = \begin{cases} \sum_n \bar{\psi}_{n\alpha}^{(-)}(x) \bar{\psi}_{n\beta}^{(+)}(x'), & t' < t, \\ \sum_n \bar{\psi}_{n\alpha}^{(+)}(x) \bar{\psi}_{n\beta}^{(-)}(x'), & t' > t, \end{cases} \quad (18.30)$$

where $\psi_n^{(+)}$ and $\psi_n^{(-)}$ are positive and negative frequency eigenfunctions of the electron in the external field.

We note that as in the case of the D -functions, we may write

¹⁾ R. Feynman, Phys. Rev. 76, 760 (1949).

$$\Delta^F = \Delta^{(1)} - i\epsilon\Delta.$$

In the future it shall be necessary to know the Fourier transforms of $\Delta^F(x)$ and $\Delta^F(x)$.
Setting

$$\left. \begin{aligned} \Delta^F(x) &= \int \Delta^F(p) e^{ipx} d^4p, \\ S_{\psi\psi}^F(x) &= \int S_{\psi\psi}^F(p) e^{ipx} d^4p \end{aligned} \right\} \quad (18.31)$$

and proceeding as we did in the derivation of (16.32), we obtain

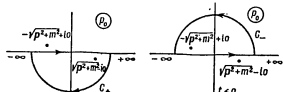
$$\Delta^F(p) = \frac{1}{(2\pi)^4} \delta_+(-p^2 - m^2) = \frac{2}{(2\pi)^4} \left\{ \pi \delta(-p^2 - m^2) - \frac{i}{p^2 + m^2} \right\}. \quad (18.31')$$

Integration over p_0 in (18.31) should be performed according to the rules given in (16.30). It is possible, as was explained in Section 16, to leave out the δ -function in the expression for δ_+ and to integrate along a contour passing above the pole $p_0 = +\sqrt{p^2 + m^2}$ and below the pole $p_0 = -\sqrt{p^2 + m^2}$, or, which is the same, not to deform the contour of integration but to shift the first pole into the lower half-plane and the second into the upper half-plane. This procedure can also be formulated differently, namely p^2 should be replaced by $p^2 - i\epsilon$, where ϵ is an infinitesimal positive quantity, and the integration over p_0 should be taken along the contour C_+ if $t > 0$, and along the contour C_- if $t < 0$ (see Fig. 5).

We note that we proceeded similarly in considering the Fourier transform of $\Delta^F(x)$.

The integration procedure can also be formulated in the following way: the mass m should be considered complex with an infinitesimal negative imaginary part,

$$m \rightarrow m - i\epsilon. \quad (18.32)$$



This procedure can be used both for $\Delta^F(x)$ and for $\Delta^F(x)$, except that in the second case it is necessary to set $m=0$.

We shall henceforth use this procedure, writing $\Delta^F(p)$ in the form

$$\Delta^F(p) = \frac{2}{i} \frac{1}{(2\pi)^4} \frac{1}{p^2 + m^2}, \quad m \rightarrow m - i\epsilon. \quad (18.33)$$

From this and from (18.29) it follows that

$$S_{\psi\psi}^F(p) = \frac{2}{i} \frac{1}{(2\pi)^4} \frac{(i\gamma_\mu p_\mu - m)\epsilon_2}{p^2 + m^2}, \quad m \rightarrow m - i\epsilon. \quad (18.34)$$

(18.31) and (18.33) can be used to show that $\Delta^F(x)$ satisfies the equation

$$(\square - m^2)\Delta^F(x) = 2i\delta(x). \quad (18.35)$$

4. Ordered Products of the Field Operators.

Let us now consider the ordered products of the electron-positron field components.

Let us write ψ and $\bar{\psi}$ in the form

$$\psi(x) = u(x) + \bar{v}(x), \quad \bar{\psi}(x) = \bar{u}(x) + v(x), \quad (18.36)$$

where $u(x)$ and $\bar{v}(x)$ are those parts of expression (17.8) for ψ which contain the positive and negative frequencies, respectively (that is, the terms with $e^{i\frac{1}{2}p\cdot x}$ and those with $e^{-i\frac{1}{2}p\cdot x}$); $v(x)$ and $\bar{u}(x)$ play the same roles for the operator $\bar{\psi}$. Using (17.13) we may say that $u(x)$ and $v(x)$ are electron and positron annihilation operators, and that $\bar{u}(x)$ and $\bar{v}(x)$ are electron and positron emission operators.

In the ordered products all the annihilation operators should be on the right and all the creation operators should be on the left.

Ordering the operators in this way, as we shall see later (see Section 22), is very convenient in determining matrix elements of products of operators, since in this case the annihilation operators annihilate only those particles which are in the original state, and the emission operators create only such particles which are found in the final state.

The ordered product of operators $\psi_1, \psi_2, \dots$ which we shall denote by $N(\psi_1\psi_2\dots)$, is given by the distributive law

$$N(\psi_1\psi_2(u + \bar{v})\psi_3\dots) = N(\psi_1\psi_2u\psi_3\dots) + N(\psi_1\psi_2\bar{v}\psi_3\dots) \quad (18.37)$$

and the rule according to which

$$N(UV\dots Z) = \delta_{XY}\dots W, \quad (18.38)$$

where each of the operators $U, V, \dots, Z$ is either an emission operator ($\bar{u}, \bar{v}$) or an absorption operator (u, v); on the right side of (18.38) these operators are ordered so that the emission operators are on the left and the absorption operators on the right, and δ_P is $+1$ or -1 depending on whether the permutation $UV\dots Z \rightarrow XY\dots W$ is even or odd.

Some examples of N -products are

$$\left. \begin{aligned} N(\bar{\psi}(x) \bar{\psi}(y)) &= N[(u(x) + \bar{v}(x))(v(y) + \bar{u}(y))] = \\ &= -\bar{u}(y)u(x) + u(x)v(y) + \bar{v}(x)\bar{u}(y) + \bar{v}(x)v(y), \\ N(\bar{\psi}(y) \bar{\psi}(x)) &= \bar{u}(y)u(x) + v(y)u(x) + \bar{u}(y)\bar{v}(x) - \bar{v}(x)v(y), \\ N(\bar{\psi}(x) \bar{\psi}(y)) &= \bar{\psi}(x)\bar{\psi}(y), \\ N(\bar{\psi}(x) \bar{\psi}(y)) &= \bar{\psi}(x)\bar{\psi}(y). \end{aligned} \right\} \quad (18.39)$$

We note that

$$N(\bar{\psi}(x) \bar{\psi}(y)) + N(\bar{\psi}(y) \bar{\psi}(x)) = 0. \quad (18.39')$$

It is easy to see that expression (18.16) for the current can be written in the form of the N -product

$$j_\mu = ieN(\bar{\psi}\gamma_\mu\psi) = ie(\gamma_\mu)_{\alpha\beta}N(\bar{\psi}_\alpha\psi_\beta). \quad (18.40)$$

In fact,

$$\begin{aligned} j_\mu &= \frac{ie}{2} (\gamma_\mu)_{\alpha\beta} (\bar{\psi}_\alpha\psi_\beta - \bar{\psi}_\beta\psi_\alpha) = \frac{ie}{2} (\gamma_\mu)_{\alpha\beta} [(\bar{u}_\alpha + v_\alpha)(u_\beta + \bar{v}_\beta) - (u_\beta + \bar{v}_\beta)(\bar{u}_\alpha + v_\alpha)] = \\ &= \frac{ie}{2} (\gamma_\mu)_{\alpha\beta} [\bar{u}_\alpha u_\beta + \bar{u}_\alpha \bar{v}_\beta + v_\alpha u_\beta + v_\alpha \bar{v}_\beta - u_\beta \bar{u}_\alpha - u_\beta v_\alpha - \bar{v}_\beta \bar{u}_\alpha - \bar{v}_\beta v_\alpha]. \end{aligned}$$

But

$$\begin{aligned} \bar{u}_\alpha \bar{v}_\beta + \bar{v}_\beta \bar{u}_\alpha &= 0, \quad v_\alpha u_\beta + u_\beta v_\alpha = 0, \\ \bar{u}_\alpha u_\beta + u_\beta \bar{u}_\alpha &= v_\alpha \bar{v}_\beta + \bar{v}_\beta v_\alpha. \end{aligned}$$

Therefore,

$$j_\mu = ie(\gamma_\mu)_{\alpha\beta} (\bar{u}_\alpha u_\beta + \bar{u}_\alpha \bar{v}_\beta + v_\alpha u_\beta - \bar{v}_\beta v_\alpha).$$

Equation (18.40) gives the same expression:

$$j_\mu = ie(\gamma_\mu)_{\alpha\beta} N(\bar{\psi}_\alpha\psi_\beta) = ie(\gamma_\mu)_{\alpha\beta} (\bar{u}_\alpha u_\beta + \bar{u}_\alpha \bar{v}_\beta + v_\alpha u_\beta - \bar{v}_\beta v_\alpha).$$

The ordered product of operators can be simply expressed in terms of their chronological product. We shall show, first, that

$$e(t, t') P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) - N(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) = \frac{1}{2} S_{\alpha\beta}^{\pm}(x' - x). \quad (18.41)$$

Using (17.8) and the quantum condition (17.7), we obtain

$$\begin{aligned} e(t, t') P(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) - N(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) &= \\ &= \begin{cases} \frac{1}{V} \sum_p \sum_{r=1}^{\infty} \bar{u}_\alpha^p \bar{u}_\beta^p e^{-ip(x-x')} & t > t', \\ -\frac{1}{V} \sum_p \sum_{r=1}^{\infty} \bar{u}_\alpha^p \bar{u}_\beta^p e^{-ip(x-x')} & t < t'. \end{cases} \end{aligned} \quad (18.41')$$

From the definition of $S_{\alpha\beta}^{\pm}$ [see (18.27), (18.28)], we obtain (18.41).

Later, in addition to the chronological product, we shall need the so-called T -product, which is defined in the following way:⁹⁾

$$T(\bar{\psi}_1 \bar{\psi}_2 \dots) = \bar{\psi}_p \bar{\psi}_i \bar{\psi}_k \dots \quad (18.42)$$

where the operators $\bar{\psi}_1, \bar{\psi}_2, \dots$ are chronologically ordered, so that $t_1 > t_2 > \dots$, and δ_P is $+1$ or -1 depending on whether the permutation $\bar{\psi}_1 \bar{\psi}_2 \dots \rightarrow \bar{\psi}_i \bar{\psi}_1 \dots$ is even or odd (among the $\bar{\psi}_1 \bar{\psi}_2 \dots$ there may be both operators of the type $\bar{\psi}$, and those of the type $\bar{\psi}$).

⁹⁾ G. Wick, Phys. Rev. 80, 269 (1950).

Obviously, the chronological product differs from the T -product only by the factor δ_p :

$$T(\dots) = \delta_p P(\dots). \quad (18.43)$$

Equation (18.41) can be rewritten

$$T(\bar{\psi}_\alpha(x) \psi_\beta(x')) - N(\bar{\psi}_\alpha(x) \psi_\beta(x')) = \frac{1}{2} S_{\beta\alpha}^T(x' - x). \quad (18.44)$$

Similarly, it can be shown that

$$T(\psi_\beta(x') \bar{\psi}_\alpha(x)) - N(\psi_\beta(x') \bar{\psi}_\alpha(x)) = -\frac{1}{2} S_{\beta\alpha}^T(x' - x), \quad (18.45)$$

$$T(\psi_\beta(x') \phi_\alpha(x)) - N(\psi_\beta(x') \phi_\alpha(x)) = 0, \quad (18.46)$$

$$T(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) - N(\bar{\psi}_\alpha(x) \bar{\psi}_\beta(x')) = 0. \quad (18.47)$$

5. Representations of the Singular Functions.

In conclusion to this section we shall consider various representations of the singular functions $\Delta(x)$ and $\Delta^{(1)}(x)$, as well as the related functions $\Delta^R(x)$ and $\Delta^A(x)$ (see below).

The functions $\Delta(x)$ and $\Delta^{(1)}(x)$ can be simply expressed in terms of the cylindrical functions $J_1(z)$, $N_1(z)$ and $K_1(z)$:

$$\Delta(x) = \frac{1}{4\pi} \delta(\lambda) - \frac{m^2}{8\pi} \frac{1}{m\sqrt{\lambda}} J_1(m\sqrt{\lambda}) \frac{1 + \operatorname{sgn} \lambda}{2},$$

$$\Delta^{(1)}(x) = \begin{cases} \frac{m^2}{4\pi} \frac{N_1(m\sqrt{\lambda})}{m\sqrt{\lambda}}, & \lambda > 0, \\ \frac{m^2}{2\pi} \frac{K_1(m\sqrt{|\lambda|})}{m\sqrt{|\lambda|}}, & \lambda < 0, \end{cases} \quad (18.48)$$

where $\lambda = t^2 - \mathbf{x}^2$.

These relations will be proved by introducing the function

$$\chi(x, t) = \Delta^{(1)}(x) + i\Delta(x) = \frac{1}{(2\pi)^3} \int e^{i p x} \frac{e^{i p t}}{p} d^3 p.$$

Since

$$\chi(x, -t) = \Delta^{(1)}(x) - i\Delta(x),$$

we have

$$\left. \begin{aligned} \Delta^{(1)}(x) &= \frac{1}{2} [\chi(x, t) + \chi(x, -t)], \\ \Delta(x) &= \frac{1}{2i} [\chi(x, t) - \chi(x, -t)]. \end{aligned} \right\} \quad (18.49)$$

If, in the integral defining $\chi(\underline{x}, t)$, we integrate over the angle between $\underline{p}$ and $\underline{x}$, we obtain

$$\chi(x, t) = -\frac{1}{(2\pi)^3} \frac{\partial}{\partial r} U(r, t), \quad r = |\mathbf{x}|,$$

where

$$U(r, t) = \int_{-\infty}^{\infty} \frac{dp}{\sqrt{p^2 + m^2}} e^{i(p r + \sqrt{p^2 + m^2} t)} = \int_{-\infty}^{\infty} e^{i m (r \sinh t + k \cosh t)} dk$$

(We have performed the substitution $p = m \sinh \xi$. The last integral can be evaluated in terms of one of the following Hankel functions¹⁾).

$$H_0^{(1)}(x) = \frac{1}{i\pi} \int_{-\infty}^{\infty} e^{i x \cosh t} dt, \quad x > 0,$$

¹⁾ See Watson, Theory of Bessel Functions (Foreign Lit. Press, 1949).

$$H_0^{(0)}(x) = -\frac{1}{i\kappa} \int_{-\infty}^{\infty} e^{-i\kappa \cosh t} dt, \quad x > 0,$$

$$H_0^{(0)}(ix) = \frac{1}{\kappa} \int_{-\infty}^{\infty} e^{\pm i\kappa \sinh t} dt, \quad x > 0.$$

Indeed, replacing t by $t + \alpha$, $U(t, \pm)$ can be written

$$U(r, t) = \int_{-\infty}^{\infty} e^{i\kappa t} (A \sinh t + B \cosh t) d\kappa,$$

where

$$A = r \cosh \alpha - i \sinh \alpha, \\ B = r \sinh \alpha + i \cosh \alpha, \quad A^2 - B^2 = r^2 - 1.$$

Let us now choose α so that one of the coefficients A or B vanishes. There are then three possibilities:

- 1) $t > r$, in which case we set $\tanh \alpha = -\frac{r}{t}$, $A = 0$, $B = +\sqrt{t^2 - r^2}$;
- 2) $-r < t < r$, in which case we set $\tanh \alpha = -\frac{t}{r}$, $B = 0$, $A = +\sqrt{r^2 - t^2}$;
- 3) $t < -r$, in which case we set $\tanh \alpha = -\frac{r}{t}$, $A = 0$, $B = -\sqrt{t^2 - r^2}$.

We then arrive at the following result:

$$U(r, t) = \int_{-\infty}^{\infty} e^{i\kappa t} (\sinh t + \cosh t) d\kappa = \begin{cases} i\pi H_0^{(0)}(m\sqrt{t^2 - r^2}), & t > r, \\ \pi H_0^{(0)}(im\sqrt{r^2 - t^2}), & |t| < r, \\ -i\pi H_0^{(0)}(m\sqrt{t^2 - r^2}), & t < -r. \end{cases} \quad (18.49)$$

The functions $\Delta(x)$ and $\Delta^{(1)}(x)$ are related to $U(t, \pm)$ by

$$\Delta^{(0)}(x) = -\frac{1}{2(2\pi)^2} \frac{\partial}{\partial r} \{U(r, t) + U(r, -t)\}, \\ \Delta^{(1)}(x) = -\frac{1}{2i(2\pi)^2} \frac{\partial}{\partial r} \{U(r, t) - U(r, -t)\}, \quad (18.49')$$

where, according to (18.49),

$$\frac{1}{2} \{U(r, t) + U(r, -t)\} = \begin{cases} -\pi N_0(m\sqrt{\lambda}), & \lambda > 0, \\ \pi H_0^{(0)}(im\sqrt{|\lambda|}), & \lambda < 0; \end{cases} \\ \frac{1}{2i} \{U(r, t) - U(r, -t)\} = \begin{cases} \pi J_0(m\sqrt{\lambda}) \operatorname{sgn} t, & \lambda > 0, \\ 0, & \lambda < 0 \end{cases}$$

and $\lambda = t^2 - r^2$. We have used the well-known formulas

$$\frac{i}{2} [H_0^{(0)}(x) - H_0^{(0)}(x)] = N_0(x), \\ \frac{1}{2} [H_0^{(0)}(x) + H_0^{(0)}(x)] = J_0(x).$$

Differentiating these expressions by t and making use of (18.10'), we obtain (18.48).

From (18.48) it follows that for small values of λ , the functions $\Delta(x)$ and $\Delta^{(1)}(x)$ can be written⁴⁾

$$\Delta(x) = \frac{1}{4\pi} \left\{ \delta(\lambda) + \frac{m^2}{4} \theta(\lambda) \right\}, \\ \Delta^{(1)}(x) = \frac{1}{4\pi} \left\{ -\frac{2}{\lambda} + m^2 \ln \frac{1}{2} \sqrt{|\lambda|} m^2 - \frac{m^2}{2} \right\}, \quad (18.50)$$

where

$$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases}$$

The functions $\Delta(x)$ and $\Delta^{(1)}(x)$ can also be written in the form of contour integrals, namely

$$\Delta(x) = -\frac{1}{(2\pi)^2} \oint \frac{e^{ipx}}{p^2 + m^2} d^4p, \\ \Delta^{(1)}(x) = \frac{1}{i(2\pi)^2} \oint \frac{e^{ipx}}{p^2 + m^2} d^4p, \quad (18.51)$$

⁴⁾ See Pauli and Villars, Rev. Mod. Phys. 21, 434 (1949); this has been translated into Russian and appears in Atomic Electron Level Shifts (Foreign Lit. Press, 1950).

where the integration over $\mathcal{D}_B$ is performed along the contours $\mathcal{C}$ and $\mathcal{C}_1$ shown in Fig. 6.

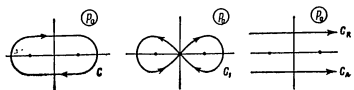


Fig. 6

We shall later also need the function Δ^R defined by the contour integral

$$\Delta^R(x) = \frac{1}{(2\pi)^4} \int_{\mathcal{C}_R} \frac{e^{ipx}}{p^2 + m^2} d^4p, \quad (18.52)$$

where integration over $\mathcal{D}_B$ is performed along the contour $\mathcal{C}_R$ shown in Fig. 6.

It is easily shown that the function $\Delta^R(x)$ is a solution of the equation

$$(\square - m^2) \Delta^R(x) = -\delta(x), \quad (18.53)$$

which vanishes for $t < 0$ and is equal to $-\Delta(x)$ for $t > 0$:

$$\Delta^R(x) = \begin{cases} 0, & t < 0 \\ \Delta(x), & t > 0. \end{cases} \quad (18.53')$$

In Section 21 we shall show that $\Delta^R(x)$ is simply related to the Green's function for the Dirac equation.

We note that the function

$$\Delta^A(x) = \frac{1}{(2\pi)^4} \int_{\mathcal{C}_A} \frac{e^{ipx}}{p^2 + m^2} d^4p,$$

where the contour of integration $\mathcal{C}_A$ over $\mathcal{D}_B$ is shown in Fig. 6, is a solution to the equation;

$$(\square - m^2) \Delta^A(x) = -\delta(x),$$

and that

$$\Delta^A(x) = \begin{cases} -\Delta(x), & t < 0, \\ 0, & t > 0. \end{cases} \quad (18.53'')$$

CHAPTER IV

FUNDAMENTAL EQUATIONS OF QUANTUM ELECTRODYNAMICS

§ 19. Interaction Between the Electron-Positron and the Electromagnetic Fields.

1. Fundamental Equations of the Interacting Fields.

We have thus far considered the free electron-positron and electromagnetic fields. We shall now go on to an examination of the interaction between these fields.

Let us first recall the fundamental relations of the classical theory of the interaction between the electron and the electromagnetic field.

In the absence of charges the field potential A_μ satisfies the homogeneous equation (15.7). If there are charges present, then the potential due to these charges is a solution of the inhomogeneous wave equation

$$\square A_\mu = -j_\mu, \quad (19.1)$$

where j_μ ($\rho \mathbf{v}$, $i\rho$) is the four-vector current density (ρ is the charge density, and $\mathbf{v}$ is the velocity of the charges).¹⁾ This equation makes it possible to find the field if the charges and their motion is known.

A particle with a charge e in the electromagnetic field is subjected to a force $\mathbf{f} = e(\mathbf{E} + [\mathbf{v} \mathbf{H}])$ (where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field strengths), so that its equation of motion becomes

$$\frac{d}{dt} \mathbf{p} = e(\mathbf{E} + [\mathbf{v} \mathbf{H}]),$$

where $\mathbf{p}$ is the momentum of the particle. To this equation corresponds the Hamiltonian

$$H = \sqrt{(\mathbf{p} - e\mathbf{A})^2 + m^2} + eA_0,$$

where $\mathbf{p}$ is the generalized momentum of the particle (defined as the derivative of the Lagrangian with respect to the velocity).

¹⁾ Here and henceforth we shall use Heaviside units.

Comparing this expression with the Hamiltonian for a free particle, namely $H = \sqrt{\mathbf{p}^2 + m^2}$, we see that the Hamiltonian of a particle in a field can be obtained from the free-particle Hamiltonian by replacing its four-vector momentum p_μ by $p_\mu - eA_\mu$ according to

$$p_\mu \rightarrow p_\mu - eA_\mu, \quad (19.2)$$

where p_μ is the generalized four-dimensional momentum of the particle.

Let us now go on to the quantum theory of the interaction between an electron and the electromagnetic field.

It is known that the classical generalized momentum p_μ is replaced in quantum mechanics by the differential operator $\frac{1}{i} \frac{\partial}{\partial x_\mu}$. On the basis of (19.2) we postulate that the wave equation of the electron in a given electromagnetic field A_μ can be obtained from the free-electron wave equation by performing the substitution

$$\frac{1}{i} \frac{\partial}{\partial x_\mu} \rightarrow \frac{1}{i} \frac{\partial}{\partial x_\mu} - eA_\mu. \quad (19.3)$$

Applying this concept to the Dirac equation (7.30), we have obtained Equation (11.1), namely

$$\left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ieA_\mu \right) + m \right] \psi = 0, \quad (19.4)$$

which is the wave equation of an electron in a given electromagnetic field.

If we now construct the four-vector

$$j_\mu = ie\bar{\psi} \gamma_\mu \psi, \quad (19.5)$$

where $\bar{\psi} = \psi^\dagger \gamma_4$, then using (19.4) and the equation

$$\bar{\psi} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) - m \right] = 0, \quad (19.6)$$

satisfied by $\bar{\psi}$, it is easy to see that the vector j_μ satisfies the continuity equation

$$\frac{\partial j_\mu}{\partial x_\mu} = 0. \quad (19.7)$$

This vector, like (17.1) in the case of the free electron, shall be considered the four-vector current density of the electron in the state ψ .

The current density vector $\underline{j}_\mu$ can be used to find the electromagnetic field due to the electron in the state ψ ; this field satisfies the wave equation (19.1), in which the current is given by (19.5).

Considering Equation (19.4) as the wave equation of an individual electron in a given electromagnetic field, and Equation (19.1) as the wave equation of the classical electromagnetic field due to a given electron current, we can solve many problems; among these are the scattering and energy eigenvalues of an electron in a given magnetic field, as well as radiation, absorption, and scattering of light when an electron goes from one state to another (compare Section 35). This approach, however, is not sufficiently general and cannot be considered satisfactory; indeed, the particle nature of the electromagnetic field and the possibility that the number of electrons may not be conserved (electron-positron pair creation and annihilation in the electromagnetic field) are not included.

In order to formulate the correct quantum mechanical theory of the interaction between the electron-positron and electromagnetic fields, we shall start from the fact that both fundamental aspects, that is, both the corpuscular and the wave properties, can be described with the aid of quantum field theory, as has been shown in Chapter III; in this theory the electromagnetic field potential A_μ and the electron wave function ψ are not considered ordinary c-numbers, but operators in occupation-number space which, as functions of space and time, satisfy definite wave equations. In the case of the free, noninteracting fields these wave equations were D'Alembert's equation (15.7) for the electromagnetic potential and the Dirac equation (7.30) for the electron-positron field. We shall now assume that the interacting electromagnetic and electron-positron fields can be described with the aid of operators $\underline{A}_\mu$ and ψ which satisfy Equations (19.1), (19.4), and (19.6) as functions of space and time.

As for the current $\underline{j}_\mu$ on the right side of (19.1), we may define it according to (19.5) or, since ψ and $\bar{\psi}$ are now considered operators, according to (18.23):

$$j_\mu = \frac{ie}{2} (\bar{\psi} \gamma_\mu \psi - \bar{\psi}' \gamma_\mu \psi') = \frac{ie}{2} (\bar{\psi}, \gamma_\mu \psi), \quad (19.8)$$

where ψ' and $\bar{\psi}'$ are given by

$$\psi' = \bar{\psi} C', \quad \bar{\psi}' = C'^{-1} \psi \quad (19.8')$$

and satisfy the equation¹⁾

$$\left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ie A_\mu \right) + m \right] \psi' = 0, \quad \bar{\psi}' \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ie A_\mu \right) - m \right] = 0. \quad (19.9)$$

¹⁾ The equation for ψ' differs from that for ψ only in the sign of the charge.

It can be shown on the basis of these equations and (19.4) and (19.6), that the divergence of the vector $\underline{j}_\mu$ defined by (19.8), like that of $\underline{j}_\mu$ defined by (19.5), vanishes.

Thus, we shall start from the following fundamental equations of the electron-positron and electromagnetic fields:

$$\left. \begin{aligned} (\gamma_\mu \frac{\partial}{\partial x_\mu} + m) \psi &= ie \gamma_\mu A_\mu \psi, \\ \bar{\psi} (\gamma_\mu \frac{\partial}{\partial x_\mu} - m) &= -ie \bar{\psi} \gamma_\mu A_\mu, \\ \frac{\partial^2}{\partial x_\mu^2} A_\mu &= -j_\mu = -\frac{ie}{2} (\bar{\psi}, \gamma_\mu \psi). \end{aligned} \right\} \quad (19.10)$$

These determine the time and space dependence of the field operators.

These operators act on the wave function (state vector) Φ_0 of our system, which is now the system of interacting fields.

The wave function Φ_0 shall be considered constant. The change of state of the system is accounted for by the fact that the interacting field operators differ from the free field operators. In other words we are dealing with the so-called Heisenberg representation of quantum mechanics, in which the operators vary, but the wave function remains constant.

We shall not, however, allow all wave functions Φ_0 . Just as in the case of the free electromagnetic field, only such state vectors Φ_0 are permissible which satisfy the subsidiary condition

$$(\Phi_0, \frac{\partial A_\mu}{\partial x_\mu} \Phi_0) = 0. \quad (19.11)$$

From this and from the third of Equations (19.10) it is easy to obtain the second pair of Maxwell's equations, which now means that the expectation value of the operator $\frac{\partial}{\partial x_\mu} F_{\mu\nu} + j_\nu$ in the state Φ_0 vanishes, or in other words

$$(\Phi_0, \left(\frac{\partial}{\partial x_\mu} F_{\mu\nu} + j_\nu \right) \Phi_0) = 0, \quad (19.12)$$

where $F_{\mu\nu}$ is the field tensor operator

$$F_{\mu\nu} = \frac{\partial}{\partial x_\mu} A_\nu - \frac{\partial}{\partial x_\nu} A_\mu.$$

To Equations (19.10) and (19.11) we must add the quantum conditions for the field operators. As will be shown below, it is not necessary to know the quantum conditions in general, but it is sufficient to know them only for the free, noninteracting fields which we have already studied.

2. Variational Principle.

Before going on to a consideration of the methods for solving the quantum field equations (19.10), we shall show that these equations can be obtained from a variational principle

$$\delta \int L d^4x = 0,$$

in which the Lagrangian L is

$$L = -\frac{1}{2} \frac{\partial A_\mu}{\partial x_\nu} \frac{\partial A_\mu}{\partial x_\nu} - \frac{1}{2} \bar{\psi} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ieA_\mu \right) + m \right] \psi - \frac{1}{2} \bar{\psi} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) + m \right] \psi' + j_\mu A_\mu \quad (19.13)$$

and the independent variables for the variation are $A_\mu, \psi, \bar{\psi}$ (or $A_\mu, \psi', \bar{\psi}'$).

Indeed, according to (18.22)

$$\bar{\psi} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) + m \right] \psi' = \bar{\psi}' \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) - m \right] \bar{\psi},$$

and therefore, performing the variation and dropping the unessential terms which can be written as divergences, we obtain

$$\delta L = \frac{1}{2} \delta A_\mu (\square A_\mu + j_\mu) + \frac{1}{2} (\square A_\mu + j_\mu) \delta A_\mu - \frac{1}{2} \delta \bar{\psi} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ieA_\mu \right) + m \right] \psi + \frac{1}{2} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ieA_\mu \right) + m \right] \bar{\psi} \delta \psi - \frac{1}{2} \delta \bar{\psi}' \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) - m \right] \bar{\psi}' + \frac{1}{2} \left[\gamma_\mu \left(\frac{\partial}{\partial x_\mu} + ieA_\mu \right) - m \right] \bar{\psi}' \delta \bar{\psi} = 0, \quad (19.14)$$

from which we get (19.10).

If, in varying L , we had considered not ψ and $\bar{\psi}$, but ψ' and $\bar{\psi}'$ to be the independent variables, we would have obtained Equations (19.9) for the charge conjugate operators $\bar{\psi}$ and $\bar{\psi}'$ instead of the first two equations of (19.10).

We shall show that the Lagrangian is invariant under three kinds of transformations; Lorentz transformations, gauge transformations, and charge conjugation.

We needn't go into the Lorentz transformations, since this invariance is obvious from the form of the Lagrangian.

A gauge transformation is given by the substitutions

$$\left. \begin{aligned} A_\mu &\rightarrow A_\mu - \frac{\partial \eta}{\partial x_\mu}, \\ \psi &\rightarrow e^{-ie\eta} \psi, \\ \psi' &\rightarrow e^{ie\eta} \psi', \end{aligned} \right\} \quad (19.15)$$

where η is a scalar function of x . Inserting (19.15) into (19.13), we obtain the following addition to the Lagrangian:

$$\frac{\partial}{\partial x_\mu} \left(\frac{1}{2} \left(A_\mu \frac{\partial^2 \eta}{\partial x_\mu \partial x_\nu} + \frac{\partial^2 \eta}{\partial x_\mu \partial x_\nu} A_\mu \right) - \frac{1}{2} \left(A_\mu \frac{\partial}{\partial x_\nu} \frac{\partial^2 \eta}{\partial x_\mu \partial x_\nu} + \frac{\partial}{\partial x_\nu} \frac{\partial^2 \eta}{\partial x_\mu \partial x_\nu} A_\mu \right) - \frac{1}{2} \left(\frac{\partial^2 \eta}{\partial x_\mu \partial x_\nu} \right)^2 \right), \quad (19.16)$$

The first of these terms is a divergence and, therefore, is not significant, and the third is independent of the fields; the second term vanishes if

$$\square \eta = 0. \quad (19.17)$$

Thus, the Lagrangian is invariant under gauge transformations if the scalar function η satisfies the wave equation, but is otherwise arbitrary. We note that the same class of functions η was used for the gauge invariance of the free electromagnetic field (see Section 15).

Charge conjugation is the operation of going over from the field operators $\bar{\psi}, \bar{\psi}'$ to the charge conjugate operators $\bar{\psi}', \bar{\psi}$ and simultaneously changing the sign of the potential⁹.

$$\bar{\psi} \rightarrow \bar{\psi}', \quad \bar{\psi}' \rightarrow \bar{\psi}, \quad A_\mu \rightarrow -A_\mu. \quad (19.18)$$

The invariance of the Lagrangian under such transformation can be seen directly from the expression for L . This invariance is a mathematical formulation of the symmetry of the theory with respect to the sign of the charge, i.e. with respect to exchanging electrons and positrons.

⁹ Changing the sign of the potential is equivalent to changing the sign of the charge; $e \rightarrow -e$.

3. Charge Parity.

To the transformation of the field operators according to (19.18) there corresponds a transformation of the wave function of the electron and photon system. We shall write this transformation

$$\Phi' = \Lambda \Phi. \quad (19.19)$$

The transformations (19.18) and (19.19) are related, as is well known, in the following way:

$$\left. \begin{aligned} \Phi &\rightarrow \Phi' = \Lambda \Phi \Lambda^{-1}, \\ \Lambda &\rightarrow \Lambda' = \Lambda \Lambda \Lambda^{-1}. \end{aligned} \right\} \quad (19.20)$$

Comparing (19.20) with (19.8'), we obtain

$$\left. \begin{aligned} \Lambda \psi \Lambda^{-1} &= \bar{\psi} C', \\ \Lambda \bar{\psi} \Lambda^{-1} &= C'^{-1} \psi, \\ \Lambda \hat{a}_\mu \Lambda^{-1} &= -\hat{a}_\mu. \end{aligned} \right\} \quad (19.21)$$

From (19.21) we can obtain an explicit expression for the operator Λ . To do this, let us write it in the form of a product of operators acting on the photon and electron variables separately, so that

$$\Lambda = \Lambda_\gamma \Lambda_e.$$

The operator Λ_γ , which according to (19.21) anticommutes with $\hat{a}_\mu$, can be written¹⁾

$$\Lambda_\gamma = (-1)^N, \quad (19.22)$$

where N is the number of photons.

According to (19.21) and (18.25), the operator Λ_e is related in the following way to the creation and annihilation operators for electrons and positrons:

$$\left. \begin{aligned} \Lambda_e \hat{a}_r \Lambda_e^{-1} &= \hat{b}_r, \\ \Lambda_e \hat{b}_r^\dagger \Lambda_e^{-1} &= \hat{a}_r^\dagger, \\ \Lambda_e \hat{a}_r^\dagger \Lambda_e^{-1} &= \hat{b}_r^\dagger, \\ \Lambda_e \hat{b}_r \Lambda_e^{-1} &= \hat{a}_r. \end{aligned} \right\} \quad (19.23)$$

Equations (19.23) are satisfied if Λ_e is²⁾

$$\Lambda_e = \prod_r (1 - \hat{a}_r^\dagger \hat{a}_r - \hat{b}_r^\dagger \hat{b}_r + \hat{a}_r^\dagger \hat{b}_r + \hat{b}_r^\dagger \hat{a}_r). \quad (19.24)$$

It is easy to see that when the operator Λ_e is applied to the wave function, each "unoccupied" electron state is multiplied by +1, and each state occupied by an electron is transformed to an occupied positron state.

The operator Λ commutes with the energy, momentum, and angular momentum operators of the system. But it anticommutes with the total charge operator Q and is therefore, not in general an integral of the motion. When, however, $Q = 0$, the states of the system may be eigenstates of Λ . Since if Λ is applied twice, the result is the identity operation

$$\Lambda^2 = 1,$$

the eigenvalues of Λ are

$$\Lambda = \pm 1.$$

Thus, for given values of the energy, angular momentum, and parity, a neutral system can be separated into two states corresponding to different values of the "charge parity": $\Lambda = \pm 1$. From the commutation of Λ with the Hamiltonian, it follows that charge parity is conserved.

It can be seen from (19.22) that a system consisting of an even number of photons has even charge parity, and one consisting of an odd number of photons has odd charge parity. Thus, an arbitrary neutral system which is an eigenstate of the charge parity can decay either only into an even number or only into an odd number of photons. An example is the decay of the π^0 -meson³⁾ into two photons, from which it follows that the π^0 -meson is a system with even charge parity.

¹⁾ L. Wolfenstein and D. Ravenhall, Phys. Rev. 88, 279 (1952).

²⁾ In making this assertion, of course, we are leaving the framework of electrodynamics and are assuming that all interactions are invariant under charge conjugation.

³⁾ Equation (19.22) can be derived in the same way as (15.46) was derived from (15.45').

For a system consisting of one electron and one positron, (19.24) leads to the expression

$$\Lambda_e \psi(1, 2) = -\psi(2, 1), \quad (19.25)$$

where ψ is the wave function of the system in configuration (or momentum) space and the numbers 1 and 2 denote the set of coordinates (or momenta) and spin variables of the electron and positron, respectively. Equation (19.25) indicates that charge conjugation in this case corresponds to exchanging the coordinates and spins of the two particles. The minus sign in (19.25) corresponds to the anticommutation of the electron and positron operators.

It follows from (19.25) that symmetric electron-positron states have odd charge parity, and that antisymmetric states have even charge parity. If l is the relative orbital angular momentum of the particles, so that $(-1)^l$ is the parity of the state, then

$$\Lambda = \begin{cases} (-1)^l & \text{for singlet states } (s = 0). \\ (-1)^{l+1} & \text{for triplet states } (s = 1). \end{cases} \quad (19.26)$$

In particular, an electron-positron system in the 1S state ($l = 0, s = 0$) can decay only into an even number of photons, whereas a system in the 3S state ($l = 0, s = 1$) can decay only into an odd number of photons.

We note that not every neutral system need have a definite charge parity. Consider, for example, the hydrogen atom. Let us denote its wave function by ψ_H . Then

$$\Lambda \psi_H = \psi_{\bar{H}},$$

where $\psi_{\bar{H}}$ is the wave function of "antihydrogen" (consisting of an antiproton and a positron). The functions ψ_H and $\psi_{\bar{H}}$ are not simply related. The eigenfunctions of the operator Λ are

$$\psi_H \pm \psi_{\bar{H}}.$$

4. Schrodinger Equation for a System of Fields. "External" Field. "Given" Current.

As has been indicated above, the fundamental Equations (19.10) of quantum electrodynamics describe the state of the electromagnetic and electron-positron fields in the Heisenberg representation. As in ordinary quantum mechanics, one may employ also the Schrodinger representation, in which the field operators are time independent and the change of state of the system is described by the time dependence of the wave function. We shall call such a wave function $\psi(t)$. It satisfies the equation

$$i \frac{\partial \psi}{\partial t} = H \psi, \quad (19.27)$$

where

$$H = H_0 + V, \quad H_0 = H_e + H_p,$$

and H_e and H_p are the energy operators of the free electron and photon, given by (15.25) and (17.14); V is the interaction energy of the electron and the electromagnetic field.

From the form of the Lagrangian (19.13) it follows that in the Heisenberg representation

$$V = - \int j_\mu A_\mu d^3x, \quad (19.28)$$

where j is given by (19.8). In the next section we shall show how to use (19.28) to determine V in the Schrodinger representation.

Before going on to an investigation of the equations of quantum electrodynamics, we mention that in many physical problems it is convenient to separate out of the electromagnetic potential operator $A_\mu(x)$ the "external" field $A_\mu^e(x)$, and to consider it a known given c -function of space and time. If we restrict ourselves only to this field and do not take account of the additional operator $A_\mu - A_\mu^e$, the problem reduces to the investigation of the equation of motion of the electron in the given field. As an example of such a procedure, we indicate the energy eigenvalue problem, as well as the scattering of electrons by the field of the nucleus (see Sections 11-13).

In a similar way, in many problems such as the radiation from a moving electron, the electron motion is considered known; i.e., the electron field operator ψ is treated as a known c -function of space and time. When this function is known, (19.8) can be used to find the current ("given" current), and the problem reduces to solving the inhomogeneous wave equation (19.1) in which the right side of the equation is known (see Section 35).

This method, however, does not eliminate the general problem of interaction between the fields. Furthermore, this kind of treatment is not entirely rigorous, since the interaction of the electromagnetic field with the electron-positron vacuum changes the properties of the electromagnetic field, and the interaction of the electron-positron field with the electromagnetic vacuum changes the properties of the electron. As a result, the very concept of the "given external field" is not rigorous; actually this "given" field must be changed somewhat, since to it must be added the field related to the polarization of the electron-positron vacuum caused by the "external" field. In exactly the same way, the above mentioned "given" electron current must be changed, since to it we must add the current due to the interaction of the "given" current with the electromagnetic vacuum. Thus, the so-called exact solutions of problems concerning the possible energies and the scattering of an electron in some "given" field of force (for instance, the Coulomb field) are actually approximate; to the energy eigenvalues and to the scattered amplitude there must be added the so-called radiative corrections (see Chapter VIII).

Let us now return to the general field equations (19.10). Since the terms in these equations which determine the interaction between the fields are proportional to the electron charge, or to the dimensionless quantity $\sqrt{\alpha} = \frac{e}{\sqrt{4\pi\hbar c}} = \frac{1}{\sqrt{137}}$ which is much less than one, it is natural to use perturbation theory to treat these equations (see, however, paragraph 3 of Section 27). Perturbation theory can be employed both for treating Equations (19.10) which describe the fields in the Heisenberg representation, and for investigating Equation (19.27) which describes the fields in the Schrodinger representation. We shall start by describing the application of perturbation theory to the Schrodinger equation (19.27).

§ 20. Perturbation Theory. Transition from the Schroedinger Representation to the Interaction Representation.¹⁾

1. Interaction Representation.

In the Schroedinger representation the state of the fields is described by Schroedinger's equation (19.27), which can be handled with the aid of perturbation theory. To do this let us introduce the eigenfunctions of the free-field Hamiltonian H_0 , which is the sum of the free-electron Hamiltonian H_e and the free-photon Hamiltonian H_γ . Let us denote these functions by $\Psi_n(t)$. Each of the $\Psi_n(t)$ describes a definite state of the electron-and-photon system and satisfies the equation

$$i \frac{\partial \Psi_n}{\partial t} = H_0 \Psi_n = E_n \Psi_n, \quad (20.1)$$

where E_n is the energy of the system in the state Ψ_n . It is clear that

$$\Psi_n(t) = \Psi_n^{(0)} e^{-iE_n t},$$

where $\Psi_n^{(0)}$ is time dependent. We shall treat the $\Psi_n^{(0)}$ as functions of the number of electrons and photons in the various individual states ("occupation numbers"). We shall denote the individual electron states by f , and the individual photon states by q (f and q denote the full set of quantum numbers describing the states of the particles, such as the eigenvalues of the momentum and polarization), and the number of particles in states f and q by N_f and N_q . Then in occupation-number space the wave function $\Psi_n^{(0)}$ can be written

$$\Psi_n^{(0)}(N_f, N_q) = \prod_{f,q} \delta_{N_f N_f^0} \delta_{N_q N_q^0}, \quad (20.2)$$

where n indicates a definite set of numbers N_f^0, N_q^0 and f and q take on all possible values.

The functions $\Psi_n^{(0)}$ are obviously a complete orthonormal set; in fact,

$$(\Psi_n^{(0)}, \Psi_{n'}^{(0)}) = \sum_{N_f, N_q} \Psi_n^{(0)*}(N_f, N_q) \Psi_{n'}^{(0)}(N_f, N_q) = \delta_{nn'}.$$

(the equation $n = n'$ means that each separate pair of occupation numbers is equal, or $N_f^0 = N_f^0, N_q^0 = N_q^0$).

¹⁾ See also V. Berestetsky, Progr. Phys. Sci. (USSR) 46, 231 (1952).

Let us now consider the interaction between the electromagnetic and the electron-positron fields. They can clearly not be in the state described by Ψ_n , since this is not an eigenfunction of the interaction operator V . The wave function $\Psi(t)$ of the system, however, which satisfies equation (19.27), can be expanded in a series of the functions (20.2),

$$\Psi(t) = \sum_n \Phi_n(t) \Psi_n(t), \quad (20.3)$$

where the coefficients $\Phi_n(t)$ depend explicitly on the time.

If at some initial time $t = t_0$ the system is characterized by some definite distribution of electron and photon occupation numbers (let the set of these numbers be n_0), then due to the interaction this distribution will be time dependent. The quantity $|\Phi_{n_0}(t)|^2$ determines the probability that at a given time t the system is characterized by the set of occupation numbers n . Thus, finding $\Phi_n(t)$ is clearly equivalent to finding $\Psi(t)$.

Let us determine the time variation of $\Phi_n(t)$. To do this we insert (20.3) into (20.1), obtaining

$$\sum_n \left\{ i \frac{\partial \Phi_n}{\partial t} \Phi_n(t) + i \frac{\partial \Phi_n}{\partial t} \Psi_n(t) \right\} = \sum_n \{ \Phi_n(t) H_0 \Psi_n(t) + \Phi_n(t) V \Psi_n(t) \}.$$

The first terms on the left and right sides of this equation cancel identically because of (20.1). Let us take the scalar product of $\Psi_{n'}^{(0)}$ with the rest of this equation. Since the Ψ_n are an orthonormal set, we obtain

$$i \frac{\partial \Phi_n(t)}{\partial t} = \sum_{n'} (n | V | n') \Phi_{n'}(t), \quad (20.4)$$

where $(n | V | n')$ is the matrix element of the interaction energy operator V given by

$$(n | V | n') = (\Psi_n, V \Psi_{n'}) = (\Psi_n^{(0)}, V \Psi_{n'}^{(0)}) e^{-i(E_n - E_{n'})t}, \quad (20.5)$$

The interaction energy operator in this equation is given in the Schroedinger representation. But it is well-known that operators in the Heisenberg representation are nothing other than the matrices of these operators in the Schroedinger representation. Therefore, since n and n' denote states of the noninteracting fields, the set of matrix elements given by (20.5) is the Heisenberg operator $\bar{V}$ defined by (19.28) in which j and A are the current and potential operators for the free field.

Equation (20.4) can be rewritten

$$i \frac{\partial \Phi}{\partial t} = V \Phi, \quad (20.6)$$

where Φ is the set of coefficients $\Phi_n(t)$. This equation can be considered a Schrodinger equation, and Φ and V can be treated as a wave function and a Hamiltonian, respectively. When the equations are written in this way, we are in the interaction representation.⁹⁾ In the interaction representation both the operators and the wave functions depend on time. In some sense this representation is intermediate between the Heisenberg and Schrodinger representations; here the operators maintain the free-field time dependence, and the time dependence of the wave functions is related to the difference between the actual properties of the system and the properties of a system of noninteracting electrons and photons.

2. Perturbation Theory.

Let us now consider Equation (20.6). The solution of this equation may be written in the form

$$\Phi(t) = S(t) \Phi^{(0)}, \quad (20.7)$$

or in other words

$$\Phi_n(t) = \sum_{n'} \langle n | S(t) | n' \rangle \Phi_{n'}^{(0)}, \quad (20.7')$$

where $\Phi^{(0)}$ is the value of Φ at $t = t_0$, and $S(t)$ is an operator. At the initial time the system is in the state Ψ_0 , so that

$$\Phi_n^{(0)} = \delta_{n, n_0},$$

and, therefore,

$$\Phi_n(t) = \langle n | S(t) | n_0 \rangle = \langle \Psi_n, S(t) \Psi_{n_0} \rangle.$$

Inserting (20.7) into (20.6), we obtain an equation for the matrix $S(t)$,

$$i \frac{\partial}{\partial t} S(t) = V S(t), \quad (20.8)$$

with the initial condition

$$S(t_0) = 1. \quad (20.8')$$

Equation (20.8), as we shall see later, plays a most important role in all the applications of quantum electrodynamics.

To solve Equation (20.8) we shall make use of the fact already mentioned that the interaction energy between electrons and the electromagnetic field, which is proportional to the electron charge $\frac{e}{\sqrt{4\pi}} = \frac{1}{\sqrt{137}}$, can be treated as a perturbation. Therefore, Equation (20.8) can be solved by the method of successive approximations, which means that we can look for a solution for $S(t)$ in the form of a power series in e , writing

$$S(t) = \sum_{k=0}^{\infty} S_k(t), \quad (20.9)$$

where S_k is proportional to e^k . Inserting (20.9) into (20.8) and equating terms of the same order, we obtain, with the initial condition (20.8'),

$$\left. \begin{aligned} S_k(t) &= -i \int_{t_0}^t V(t') S_{k-1}(t') dt', \\ S_0(t) &= 1. \end{aligned} \right\} \quad (20.10)$$

Applying this formula successively k times, we obtain

$$S_k(t) = (-i)^k \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \dots \int_{t_0}^{t^{(k-1)}} dt^{(k)} V(t') V(t'') \dots V(t^{(k)}). \quad (20.11)$$

The multiple integral in this equation is conveniently represented in a somewhat different form. For simplicity, let us first consider the term $S_2(t)$:

$$S_2(t) = - \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' V(t') V(t''). \quad (20.12)$$

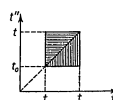


Fig. 7

In the (t', t'') plane (t' is the abscissa, and t'' is the ordinate) the region of integration in (20.12) will clearly be a triangle lying below the bisector of the angle between the coordinate axes (Fig. 7). Now interchanging the positions of t' and t'' in (20.12), we obtain

$$S_2(t) = - \int_{t_0}^t dt'' \int_{t_0}^{t''} dt' V(t'') V(t'). \quad (20.12')$$

⁹⁾ The covariant form of the interaction representation was introduced by Tomonaga and Schwinger, [S. Tomonaga, Progr. Theoret. Phys. (Japan) 1, 27 (1946); J. Schwinger, Phys. Rev. 73, 416 (1948); 74, 1439 (1948); 75, 651 (1949)].

In the (t', t'') plane (still considering t' the abscissa, and t'' the ordinate) the region of integration now becomes a triangle lying above the bisector of the coordinate angle. The integrand in expression (20.12') differs from that in (20.12) only in that the order of the factors has been changed. If $V(t')$ and $V(t'')$ commute, then both integrands are the same, and $S_2(t)$ can be written as half the integral over the whole square. This can be done in general also for noncommuting operators $V(t')$ and $V(t'')$ if we introduce the chronological operator P (see Sections 16, 18)

$$P[V(t')V(t'')] = \begin{cases} V(t')V(t''), & t' > t'', \\ V(t'')V(t'), & t' < t'', \end{cases} \quad (20.13)$$

Then

$$S_2(t) = -\frac{1}{2} \int_{t_0}^t dt' \int_{t_0}^t dt'' P[V(t')V(t'')]. \quad (20.14)$$

Similarly, it can be shown that in general $S_k(t)$ can be written (see Section 21)¹⁾

$$S_k(t) = \frac{(-i)^k}{k!} \int_{t_0}^t dt' \int_{t_0}^t dt'' \dots \int_{t_0}^t dt^{(k)} P[V(t')V(t'') \dots V(t^{(k)})], \quad (20.15)$$

where P is the chronological operator which orders the factors so that the time arguments of the operators V decrease from left to right.²⁾

§ 21. Covariant Perturbation Theory. Transition from the Heisenberg Representation to the Interaction Representation.

1. Expansion of the Interacting Field Operators in a Power Series in the Electron Charge.

Let us now go on to a consideration of the field equations (19.10) using perturbation theory. We shall attempt to find a formal solution for the field equations by expanding the field operators A and ψ in a power series in the electron charge (actually in a power series in the dimensionless quantity $\sqrt{\alpha}$)³⁾:

¹⁾ This result was obtained by Dyson [F. Dyson, Phys. Rev. 75, 486 (1949); a Russian translation can be found in the Symposium, Atomic Electron Level Shifts, p. 94].

²⁾ Having studied this section, the reader is prepared for Chapter V, which is devoted to the methods for calculating the S matrix. In Section 21 we shall derive the same results as in Section 20, using another method which is based on the equations for the field operators.

³⁾ G. Kahlen, Arkiv Fysik 2, 187, 371 (1950).

$$\left. \begin{aligned} A_\mu &= A_\mu^{(0)} + eA_\mu^{(1)} + e^2A_\mu^{(2)} + \dots, \\ \psi &= \psi^{(0)} + e\psi^{(1)} + e^2\psi^{(2)} + \dots, \\ \bar{\psi} &= \bar{\psi}^{(0)} + e\bar{\psi}^{(1)} + e^2\bar{\psi}^{(2)} + \dots \end{aligned} \right\} \quad (21.1)$$

Inserting (21.1) into (19.10) and equating the coefficients of each power of the charge to zero, we obtain the following set of equations:

$$\left. \begin{aligned} (\gamma_\mu \frac{\partial}{\partial x_\mu} + m) \psi^{(0)}(x) &= 0, \\ \square A_\mu^{(0)} &= 0; \end{aligned} \right\} \quad (21.2)$$

$$\left. \begin{aligned} (\gamma_\mu \frac{\partial}{\partial x_\mu} + m) \psi^{(n+1)}(x) &= \frac{i}{2} \sum_{m=0}^n [A_\mu^{(m)}(x), \gamma_\mu \psi^{(n-m)}(x)], \\ \square A_\mu^{(n+1)}(x) &= -\frac{i}{2} \sum_{m=0}^n [\bar{\psi}^{(m)}(x), \gamma_\mu \psi^{(n-m)}(x)] \end{aligned} \right\} \quad (21.3)$$

$(n = 0, 1, 2, \dots).$

These equations make it possible to express $\psi^{(1)}, A^{(1)}$ in terms of $\psi^{(0)}, A^{(0)}$; $\psi^{(2)}, A^{(2)}$ in terms of $\psi^{(1)}, A^{(1)}$, etc.; this means that finally we can express $\psi^{(n)}, A^{(n)}$ in terms of $\psi^{(0)}, A^{(0)}$. Therefore, in order to solve the problem completely, we must give the quantum conditions satisfied by the operators $\psi^{(0)}, A^{(0)}$. But these operators are the free-field operators, and the quantum conditions they satisfy are known. They were given in Sections 15 and 17 and can be written

$$[A_\mu^{(0)}(x), A_\nu^{(0)}(x')] = -iD(x-x')\delta_{\mu\nu}, \quad (21.4)$$

$$\left. \begin{aligned} \{\psi^{(0)}(x), \bar{\psi}^{(0)}(x')\} &= -iS_{\alpha\beta}(x-x'), \\ \{\psi^{(0)}(x), \psi^{(0)}(x')\} &= 0, \\ \{\bar{\psi}^{(0)}(x), \bar{\psi}^{(0)}(x')\} &= 0. \end{aligned} \right\} \quad (21.5)$$

The recursion equations (21.3) together with the free-field equations (21.2) and quantum conditions (21.4) and (21.5) are a formal solution of the problem.

Let us now go on to a solution of the recursion relations (21.3).

We shall first show that the particular solution of the inhomogeneous Dirac equation

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)\psi = -\chi \quad (21.6)$$

[where $\chi(x)$ is some spinor] which vanishes at $t = -\infty$, can be written in the form of an integral over all of four-space, namely

$$\psi(x) = \int S^R(x-x') \chi(x') d^4x', \quad (21.7)$$

where the Green's function $S^R(x)$ is defined by

$$S^R(x) = \left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m\right) \Delta^R(x) \quad (21.8)$$

and $\Delta^R(x)$ is given by (18.52).

Indeed, inserting (21.7) into (21.6) and making use of (21.8), we obtain

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)\psi = \int \left(\frac{\partial^2}{\partial x_\mu^2} - m^2\right) \Delta^R(x-x') \chi(x') d^4x'.$$

Since according to (18.53)

$$(\square - m^2) \Delta^R(x) = -\delta(x),$$

we may write

¹⁾ Actually this is a Green's matrix, rather than a function.

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)\psi(x) = -\int \delta(x-x') \chi(x') d^4x' = -\chi(x),$$

which shows that the function (21.7) satisfies (21.6). Since $\Delta^R(x)$ vanishes for $t < 0$, (21.7) vanishes for $t = -\infty$.

Similarly, the particular solution of the inhomogeneous equation

$$\square A_\mu(x) = -a_\mu(x), \quad (21.9)$$

which vanishes at $t = -\infty$, can be written

$$A_\mu(x) = \int D^R(x-x') a_\mu(x') d^4x', \quad (21.10)$$

where

$$D^R(x) = \Delta^R(x)_{m=0}. \quad (21.11)$$

Using (21.7) and (21.10) we can find successive solutions to the recursion relations (21.3) which vanish at $t = -\infty$.

Thus, if the interaction is "turned on" at the time $t = -\infty$, and the "free" fields $\underline{A}^{(0)}$ and $\psi^{(0)}$ are known, equations (21.3), (21.7), and (21.10) make it possible to find ψ and $\underline{A}$ at arbitrary points $\underline{x}$ ($\underline{x}$, $\underline{t}$).

It is possible, however, to construct the solution to the recursion relations (21.3) directly by using the following additional theorem.

Let $A(x)$, $B(x)$, and $H(x)$ be three operators used to construct two sets of new operators $A^{(n)}(x)$, $B^{(n)}(x)$ according to

$$\left. \begin{aligned} A^{(n)}(x) &= i^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^{(n)} [H(x^{(n)}), \dots, [H(x'), A(x)]]], \\ B^{(n)}(x) &= i^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^{(n)} [H(x^{(n)}), \dots, [H(x'), B(x)]]], \\ A^{(0)}(x) &= A(x), \quad B^{(0)}(x) = B(x), \end{aligned} \right\} \quad (21.12)$$

where $\underline{dx} = d^4x$, and the integration over $d^4x^{(n)}$ is taken over all three-space, and that over $d^4x^{(n-1)}$ is taken between the limits $-\infty$ and t^{n-1} ; then

$$(AB)^{(n)} = \sum_{k=0}^n A^{(k)} B^{(n-k)} \quad (21.13)$$

In order to prove this assertion, consider the integrand

$$R = [H(x'), \dots, [H(x'), A(x)B(x)], \dots],$$

which defines $(AB)^{(n)}$. Noting that

$$\begin{aligned} [H(x'), A(x)B(x)] &= H(x')A(x)B(x) - A(x)B(x)H(x') = \\ &= H(x')A(x)B(x) - A(x)H(x')B(x) + A(x)H(x')B(x) - A(x)B(x)H(x') = \\ &= [H(x'), A(x)]B(x) + A(x)[H(x'), B(x)] \end{aligned}$$

and applying this formula successively, we can write R in the form of 2^n terms each of which is the product of two multiple commutators

$$[H(x^m), [H(x^{m_1}), \dots, [H(x^{m_k}), A(x)]]] \times \\ \times [H(x^p), [H(x^{p_1}), \dots, [H(x^{p_{n-k}}, B(x))]]],$$

such that if H occurs in the first commutator k times, then it occurs in the second commutator $n-k$ times; the indices of the arguments of H are ordered in decreasing order:

$$m_1 > m_2 > \dots > m_k; \quad r_1 > r_2 > \dots > r_{n-k} \quad (k=0, 1, 2, \dots, n).$$

Let R_k be the sum of all those terms the first commutator of which contains k operators H . It is then clear that

$$R = \sum_{k=0}^n R_k$$

and, therefore,

$$(AB)^{(n)} = \sum_{k=0}^n I_k,$$

where

$$\begin{aligned} I_k &= \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n R_k = \\ &= \sum \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n [H(x^m), [\dots, [H(x^{m_k}), A(x)]]] \times \\ &\quad \times [H(x^p), [\dots, [H(x^{p_{n-k}}, B(x))]]] \end{aligned}$$

(the sum is taken over all terms that enter into R_k). Let us now make a substitution of variables in each integral of the last sum such that all the integrands become equal and given by

$$H(x^n), [H(x^{n-1}), \dots, [H(x^{n-k+1}), A(x)], \dots] \times \\ \times [H(x^{n-k}), [\dots, [H(x^p), B(x)], \dots]].$$

In so doing the regions of integration of different integrals will be different. Adding all these regions, we obtain

$$\begin{aligned} I_k &= \int_{-\infty}^t dx^{n-k+1} \int_{-\infty}^{t^{n-k+1}} dx^{n-k+2} \dots \int_{-\infty}^{t^{n-1}} dx^n [H(x^n), [H(x^{n-1}), \dots, [H(x^{n-k+1}), \\ &\quad A(x)], \dots]] \cdot \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-k-1}} dx^{n-k} [H(x^{n-k}), \dots, [H(x^p), \\ &\quad B(x)], \dots] = \frac{1}{k!} A^{(k)}(x) B^{(n-k)}(x), \end{aligned}$$

which proves the theorem stated in (21.13).

We shall now use the theorem of (21.13) to show that the solution of the recursion relations (21.3) can be written¹⁾

¹⁾ J. Schwinger, Phys. Rev. 74, 416; 74, 1439 (1948).

$$\left. \begin{aligned} \psi^{(n)}(x) &= i^n \int_{-\infty}^x dx' \int_{-\infty}^{x'} dx'' \dots \int_{-\infty}^{x^{n-1}} [V^{(0)}(x''), \\ &\quad [\dots, [V^{(0)}(x'), \psi^{(0)}(x)]_j \dots]], \\ A_\mu^{(n)}(x) &= i^n \int_{-\infty}^x dx' \int_{-\infty}^{x'} dx'' \dots \int_{-\infty}^{x^{n-1}} [V^{(0)}(x''), \\ &\quad [\dots, [V^{(0)}(x'), A_\mu^{(0)}(x)], \dots]], \end{aligned} \right\} \quad (21.14)$$

where

$$\left. \begin{aligned} V^{(0)}(x) &= -j_\mu^{(0)}(x) A_\mu^{(0)}(x), \\ j_\mu^{(0)}(x) &= \frac{i}{2} [\bar{\psi}^{(0)}(x), \gamma_\mu \psi^{(0)}(x)]. \end{aligned} \right\} \quad (21.15)$$

Consider the expression $f(x') [V^{(0)}(x'), \psi_\rho^{(0)}(x)]$ where $f(x')$ is some function of x' . Since the anticommutator of $\psi_\rho^{(0)}(x)$ with $\bar{\psi}_\beta^{(0)}(x')$ vanishes, and

$$[\bar{\psi}_\rho^{(0)}(x), \bar{\psi}_\beta^{(0)}(x')] = -i S_{\rho\beta}(x-x'),$$

we have

$$\begin{aligned} \int_{-\infty}^x f(x') [V^{(0)}(x'), \bar{\psi}_\rho^{(0)}(x)] dx' &= -\frac{i}{2} \int_{-\infty}^x f(x') A_\mu^{(0)}(x') [\bar{\psi}_\rho^{(0)}(x'), \\ \gamma_\mu \psi^{(0)}(x')] dx' &= \int_{-\infty}^x f(x') S_{\rho\mu}(x-x') A_\mu^{(0)}(x') (\gamma_\mu)_\rho \psi^{(0)}(x') dx'. \end{aligned}$$

From (21.8) in (18.53') it follows that

$$S^B(x) = \begin{cases} -S(x), & t > 0, \\ 0, & t < 0, \end{cases}$$

and therefore, suppressing the index ρ in the last integral, we can write it

$$\int_{-\infty}^x f(x') [V^{(0)}(x'), \bar{\psi}^{(0)}(x)] dx' = -\int_{-\infty}^x f(x') S^B(x-x') \gamma_\mu A_\mu^{(0)}(x') \bar{\psi}^{(0)}(x') dx'. \quad (21.16)$$

Let us now apply the operator $\gamma_\mu \frac{\partial}{\partial x_\mu} + m$ to Equation (21.14) for $\psi^{(n+1)}(x)$. Using (21.16) and

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) S^B(x) = (\square - m^2) S^B(x) = -\delta(x),$$

we obtain

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi^{(n+1)}(x) = i^{n+1} \int_{-\infty}^x dx' \delta(x-x') \int_{-\infty}^{x'} dx'' \dots \int_{-\infty}^{x^{n-1}} [V^{(0)}(x^{n-1}), \\ \dots, [V^{(0)}(x'), A_\mu^{(0)}(x') \gamma_\mu \bar{\psi}^{(0)}(x')], \dots]. \quad (21.17)$$

The integration over dx' reduces to replacing x' by x , after which the expression on the right differs from the operator $(A_\mu^{(0)}(x) \gamma_\mu \bar{\psi}^{(0)}(x))^{(n)}$ defined by (21.12) only by the factor i . Therefore, (21.17) can be rewritten on the basis of (21.13) in the form

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi^{(n+1)}(x) = i (A_\mu^{(0)}(x) \gamma_\mu \bar{\psi}^{(0)}(x))^{(n)} = \frac{i}{2} \sum_{k=0}^n [A_\mu^{(k)}(x), \gamma_\mu \bar{\psi}^{(n-k)}(x)].$$

We have obtained the first of the recursion relations (21.3).

Similarly it can be shown that (21.14) satisfies the second recursion relation of (21.3). Indeed, using the commutation relations (21.4) for the electromagnetic potentials, we have

$$\begin{aligned} \int_{-\infty}^x dx' f(x') [V^{(0)}(x'), A_\mu^{(0)}(x)] &= -\frac{i}{2} \int_{-\infty}^x f(x') [\bar{\psi}^{(0)}(x'), \gamma_\mu \psi^{(0)}(x')] \times \\ &\times [A_\mu^{(0)}(x'), A_\mu^{(0)}(x)] dx' = -\frac{1}{2} \int_{-\infty}^x D(x-x') f(x') [\bar{\psi}^{(0)}(x'), \gamma_\mu \psi^{(0)}(x')] dx'. \end{aligned}$$

Introducing the function $\bar{D}^R(x)$, which satisfies [see (21.11), (18.53')]]

$$D^R(x) = \begin{cases} -D(x), & t > 0, \\ 0, & t < 0, \end{cases}$$

we can rewrite the integral in the form

$$\int_{-\infty}^t f(x') [V^{(0)}(x'), A_p^{(0)}(x)] dx' = \frac{1}{2} \int_{-\infty}^{\infty} D^R(x-x') f(x') [\bar{\psi}^{(0)}(x'), \gamma_p \bar{\psi}^{(0)}(x')] dx'.$$

If we apply the operator $\square = \frac{\partial^2}{\partial x^2 \partial \nu}$ to this expression, then according to (18.53) and (21.11) we obtain

$$\square \int_{-\infty}^t f(x') [V^{(0)}(x'), A_p^{(0)}(x)] dx' = -\frac{1}{2} \int_{-\infty}^{\infty} \delta(x-x') f(x') [\bar{\psi}^{(0)}(x'), \gamma_p \bar{\psi}^{(0)}(x')] dx',$$

$$\gamma_p \bar{\psi}^{(0)}(x') dx' = -\frac{1}{2} f(x) [\bar{\psi}^{(0)}(x), \gamma_p \bar{\psi}^{(0)}(x)].$$

It then follows that if we apply the operator $\square = \frac{\partial^2}{\partial x^2 \partial \nu}$ to the expression for $A_{\mu}^{(n+1)}(x)$ given by (21.13), we obtain

$$\square A_{\mu}^{(n+1)}(x) = -\frac{1}{2} \int_{-\infty}^t dx'' \dots \int_{-\infty}^t dx^{n+1} [V^{(0)}(x^{n+1}), \dots, [V^{(0)}(x''), [\bar{\psi}^{(0)}(x), \gamma_p \bar{\psi}^{(0)}(x)]]].$$

The expression on the right differs by the factor $-\frac{1}{2}$ from the operator $(\bar{\psi}^{(n)}(x), \gamma_{\mu} \bar{\psi}^{(n)}(x))^{(n)}$ as given by (21.12). On the basis, therefore, of (21.13) we may write

$$\square A_{\mu}^{(n+1)}(x) = -\frac{1}{2} (\bar{\psi}^{(0)}(x), \gamma_p \bar{\psi}^{(0)}(x))^{(n)} = -\frac{1}{2} \sum_{k=0}^n (\bar{\psi}^{(k)}(x), \gamma_p \bar{\psi}^{(n-k)}(x)).$$

This expression is identical with the second recursion relation of (21.3).

Thus, we have shown that the solutions of Equations (21.3) can actually be represented by the integrals (21.14).

Using (21.14) and (21.1) it is easy to obtain a power series expansion for the current density in terms of the electron charge. The current density operator is given by the general formula (19.8), namely

$$j_{\mu}(x) = -\frac{ie}{2} [\bar{\psi}(x), \gamma_{\mu} \psi(x)] = \frac{ie}{2} (\gamma_{\mu})_{\alpha\beta} (\bar{\psi}_{\alpha}(x) \psi_{\beta}(x) - \psi_{\beta}(x) \bar{\psi}_{\alpha}(x)).$$

Inserting (21.1) into this expression, we obtain the power series expansion for the current density, namely

$$j_{\mu} = e j_{\mu}^{(0)} + e^2 j_{\mu}^{(1)} + \dots, \quad (21.18)$$

where

$$j_{\mu}^{(n)} = \frac{ie}{2} \sum_{k=0}^n [\bar{\psi}^{(k)}(x), \gamma_{\mu} \bar{\psi}^{(n-k)}(x)].$$

Equations (21.12) and (21.13) can be used to rewrite this expression in the form

$$j_{\mu}^{(n)}(x) = \frac{ie}{2} (\bar{\psi}^{(0)}(x), \gamma_{\mu} \bar{\psi}^{(0)}(x))^{(n)} = (j_{\mu}^{(n)}(x))^{(0)} =$$

$$= ie \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n [V^{(0)}(x^n), \dots, [V^{(0)}(x'), [V^{(0)}(x'), j_{\mu}^{(0)}(x)]]], \dots, \quad (21.19)$$

where the operator $V^{(n)}(x)$, which transforms $\bar{\psi}^{(n)}(x)$ into $\bar{\psi}^{(n)}(x)$, is defined by (21.15).

We mention that in concrete problems it is often more convenient to use (21.18) directly, and to calculate $j_{\mu}^{(n)}$ from (21.7).

2. Transition to the Interaction Representation.

We have determined the various terms of the power series expansion of the field operators in the electron charge. We shall now show that these series can be written

$$A_{\mu}(x) = S^{-1}(t) A_{\mu}^{(0)}(x) S(t), \quad (21.20)$$

$$\bar{\psi}_{\mu}(x) = S^{-1}(t) \bar{\psi}_{\mu}^{(0)}(x) S(t),$$

where the operator $S(t)$ (here t is the time coordinate of x) is given by the series

$$S(t) = \sum_{n=0}^{\infty} (-ie)^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n V^{(0)}(x') V^{(0)}(x'') \dots V^{(0)}(x^n), \quad (21.21)$$

$$V^{(0)}(x) = -j_{\mu}^{(0)}(x) A_{\mu}^{(0)}(x).$$

We shall prove (21.20) by first showing that $S(t)$ satisfies the differential equation

$$i \frac{\partial S(t)}{\partial t} = V(t) S(t), \quad (21.22)$$

where

$$V(t) = e \int V^{(0)}(x) d^3x. \quad (21.23)$$

Indeed, differentiating (21.21) by t , we have

$$\begin{aligned} \frac{\partial S(t)}{\partial t} = & -ie \int d^3x V^{(0)}(x) \sum_{n=0}^{\infty} (-ie)^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \\ & \dots \int_{-\infty}^{t^{n-1}} dx^n V^{(0)}(x^n) \dots V^{(0)}(x^n). \end{aligned}$$

Changing the variable in an obvious way, we obtain Equation (21.22), namely

$$\begin{aligned} \frac{\partial S(t)}{\partial t} = & -ie \int V^{(0)}(x) d^3x \sum_{n=0}^{\infty} (-ie)^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n V^{(0)}(x^n) \dots V^{(0)}(x^n) = \\ = & -ie \left(\int V^{(0)}(x) d^3x \right) S(t). \end{aligned}$$

It follows from (21.22) that $S(t)$, which according to (21.21) satisfies the condition $S(-\infty) = 1$, is unitary, so that

$$S(t) S^+(t) = 1 \quad (21.24)$$

(S^+ is the Hermitian conjugate of S).

Indeed, since $V(t)$ is a Hermitian operator, according to (21.22) $S^+(t)$ satisfies

$$-i \frac{\partial S^+(t)}{\partial t} = S^+(t) V(t). \quad (21.25)$$

Multiplying this equation on the right by S , and Equation (21.22) on the left by S^+ , and subtracting the first from the second we obtain

$$\frac{\partial}{\partial t} S^+(t) S(t) = 0,$$

whence follows (21.24), noting that $S(-\infty) = S^+(-\infty) = 1$.

We can now write the expression for the inverse $S^{-1}(t)$. According to (21.21) and (21.24)

$$S^{-1}(t) = S^+(t) = \sum_{n=0}^{\infty} (ie)^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n V^{(0)}(x^n) \dots V^{(0)}(x^n). \quad (21.26)$$

Expansions (21.21) and (21.26) can be used to verify (21.20). Inserting (21.21) and (21.26) into (21.20), we obtain

$$\begin{aligned} \psi(x) = & \psi^{(0)}(x) + ie \int_{-\infty}^t dx' [V^{(0)}(x') \psi^{(0)}(x) - \psi^{(0)}(x) V^{(0)}(x')] + \\ & + (ie)^2 \left[- \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' V^{(0)}(x') \psi^{(0)}(x) V^{(0)}(x'') + \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' V^{(0)}(x'') \psi^{(0)}(x) V^{(0)}(x') \right] \times \\ & \times V^{(0)}(x) \psi^{(0)}(x) + \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \psi^{(0)}(x) V^{(0)}(x') V^{(0)}(x'') + \dots \end{aligned}$$

The zeroth and first-order terms of this power series in e are $\psi^{(0)}$ and $\psi^{(1)}$ as defined by (21.14). In order to show that the quadratic terms also agree with the definition, it is sufficient to note that the integral

$$\int_{-\infty}^t \int_{-\infty}^{t'} dx' dx'' V^{(0)}(x'') \psi^{(0)}(x) V^{(0)}(x')$$

can be written

$$\int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' [V^{(0)}(x') \psi^{(0)}(x) V^{(0)}(x'') + V^{(0)}(x'') \psi^{(0)}(x) V^{(0)}(x')]$$

as can be seen by considering the region of integration over dt' . In a similar way it can be shown that the rest of the terms agree with the definition.

The identity of expressions (21.14) and (21.20) for ψ_ρ and A_μ can also be seen in the following way. Let us introduce the notation

$$F\{\psi_\rho^{(0)}(x)\} = S^{-1}(t)\psi_\rho^{(0)}(x)S(t),$$

$$G\{\psi_\rho^{(0)}(x)\} = \sum_{n=0}^{\infty} (it)^n \int_{-\infty}^t dx' \int_{-\infty}^{t'} dx'' \dots \int_{-\infty}^{t^{n-1}} dx^n [V^{(0)}(x_n), \dots, [V^{(0)}(x'), [V^{(0)}(x''), \psi_\rho^{(0)}(x)]]].$$

When $t = -\infty$, the functions F and G are equal, so that

$$F\{\psi_\rho^{(0)}(x)\} = G\{\psi_\rho^{(0)}(x)\} = \psi_\rho^{(0)}(x), \quad t = -\infty.$$

We shall now show that when $t = -\infty$, the derivatives of F and G with respect to t are also equal. Using (21.22), we find that

$$\frac{\partial F\{\psi_\rho^{(0)}(x)\}}{\partial t} = iS^{-1}(t) \left[V(t)\psi_\rho^{(0)}(x) - \psi_\rho^{(0)}(x)V(t) + \frac{1}{t} \frac{\partial \psi_\rho^{(0)}}{\partial t} \right] S(t) =$$

$$= iF\left\{ [V(t), \psi_\rho^{(0)}(x)] + \frac{1}{t} \frac{\partial \psi_\rho^{(0)}}{\partial t} \right\}.$$

In addition, it is easy to show that

$$\frac{\partial G\{\psi_\rho^{(0)}(x)\}}{\partial t} = iG\left\{ [V(t), \psi_\rho^{(0)}(x)] + \frac{1}{t} \frac{\partial \psi_\rho^{(0)}}{\partial t} \right\}.$$

Thus, $\frac{\partial F}{\partial t}$ and $\frac{\partial G}{\partial t}$ are given by $iF\{X\}$ and $iG\{X\}$, respectively, where $X = [V(t), \psi_\rho^{(0)}(x)] + \frac{1}{t} \frac{\partial \psi_\rho^{(0)}}{\partial t}$. Since F and G are equal when $t = -\infty$, then so are $\frac{\partial F}{\partial t}$ and $\frac{\partial G}{\partial t}$.

It can be shown similarly that when $t = -\infty$ all higher derivatives of F and G with respect to t are equal. It then follows that F and G are equal for all values of t , or

$$F\{\psi_\rho^{(0)}(x)\} = G\{\psi_\rho^{(0)}(x)\}.$$

Since the various terms of the power series expansion of the current density operator in e are expressed by integrals of multiple commutators whose form is the same as those in the expansions for ψ_ρ and A_μ , this operator can also be written in a way similar to (21.20), namely

$$J_\mu(x) = S^{-1}(t)j_\mu^{(0)}(x)S(t). \quad (21.27)$$

It may be said that an arbitrary operator $f(x)$, referring to the system of coupled fields, can be defined in terms of the corresponding operator $f^{(0)}(x)$ for the free fields by an expression of the form

$$f(x) = S^{-1}(t)f^{(0)}(x)S(t). \quad (21.28)$$

Therefore, $S(t)$ can be called a transformation operator, since it transforms the free field operators $\psi_\rho^{(0)}(x)$, $A_\mu^{(0)}(x)$ to the coupled field operators $\psi_\rho(x)$, $A_\mu(x)$.

Equation (21.28) may be considered a transformation connecting the field operators at some time t with the operators at the "initial" time $t = -\infty$, when the fields were free. It makes it possible to treat the interaction between the electron-positron and electromagnetic fields in a somewhat different way. Thus far we have considered the interaction between the fields to cause changes in the operators, without varying the coupled-field wave function Φ_0 . Now we can describe the interaction in a different way, considering the field operators to remain constant and equal to their values at $t = -\infty$ when the fields were free, and allowing the interaction to alter only the wave function of the system. We have already considered this method of describing the interacting fields in Section 20, calling it the interaction representation. This representation, to which we can transform both from the Schroedinger and from the Heisenberg representations, is extremely convenient for dealing with concrete problems of the coupled electron-positron and electromagnetic fields; henceforth, we shall use primarily only this representation.

As in the previous paragraph, we shall denote the wave function in the interaction representation by $\Phi(t)$. We shall prove, using (21.28), that $\Phi(t)$ satisfies Schroedinger's equation (20.6). For this purpose we note that the expectation value of any operator should obviously be the same in the Heisenberg and the interaction representations. In the Heisenberg representation, in which the form of the operator changes and the wave function remains constant, the expectation value of the operator $f(x)$ is given by

$$\langle f(x) \rangle = \langle \Phi_0, f(x)\Phi_0 \rangle. \quad (21.29)$$

The same quantity in the interaction representation, in which the wave function of the system changes and the field operators maintain their constant free-field form, is given by

$$\langle f(x) \rangle = \langle \Phi(t), f^{(0)}(x)\Phi(t) \rangle. \quad (21.30)$$

Using (21.28) and setting (21.29) equal to (21.30), we obtain

$$(\Phi_0, S^{-1}(t) i^{(0)}(x) S(t) \Phi_0) = (\Phi(t), i^{(0)}(x) \Phi(t)),$$

whence it follows immediately that

$$\Phi(t) = S(t) \Phi_0 = S(t) \Phi(-\infty). \quad (21.31)$$

Noting that $S(t)$ satisfies (21.22), we obtain Schrodinger's equation for $\Phi(t)$, namely

$$i \frac{\partial \Phi(t)}{\partial t} = V(t) \Phi(t), \quad (21.32)$$

where, as before,

$$V(t) = -e \int j_k^{(0)}(x) A_k^{(0)}(x) d^3x. \quad (21.33)$$

Equation (21.33) has already been obtained in Section 20. We see that the transformation operator $S(t)$ introduced in (21.20) is equal to that defined by (20.7).

Let us go on to a solution of Equation (21.32). The formal solution of Schrodinger's equation is usually written

$$\Phi(t) = e^{-i \int_{-\infty}^t V(t) dt} \Phi(-\infty). \quad (21.34)$$

This expression is correct if the operator $V(t)$ for a given time commutes with the same operator for some other time. If two operators A and B do not commute, then $e^{A+B} \neq e^A e^B$. Therefore, Equation (21.34) should be applied with caution if the operators $V(t_1), V(t_2), \dots$ do not commute. To indicate how to proceed in this case, we note that (21.32) can always be written

$$\Phi(t) = \lim_{\tau_n \rightarrow \tau_{n+1}} \prod_{n=0}^{\infty} \left[1 - i \int_{\tau_n}^{\tau_{n+1}} V(t) dt \right] \Phi(-\infty), \quad (21.35)$$

where the τ_n are times which approach $-\infty$ as n increases, so that

$$\tau_0 = t > \tau_1 > \tau_2 \dots$$

This expression shows that $\Phi(t)$ is first acted upon by $V(t)$ when $t = -\infty$, and then by values of this operator at later times. Therefore, expression (21.34) can be used if we are careful about the correct order of $V(t)$ for different times; earlier values of this operator should act first. To express this fact, we shall apply the chronological operator P , already discussed in Sections 16, 18, and 20, to the right side of (21.34), obtaining

$$\Phi(t) = P \left(e^{-i \int_{-\infty}^t V(t) dt} \right) \Phi(-\infty). \quad (21.36)$$

In this form the expression is always valid. The exponential appearing here can be expanded in a power series, bearing in mind the correct order of the operators. This series can be written

$$\begin{aligned} \Phi(t) &= S(t) \Phi(-\infty) = P \left(e^{-i \int_{-\infty}^t V(t) dt} \right) \Phi(-\infty) = \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^t dt_1 \dots \int_{-\infty}^t dt_n P \{ V(t_1) V(t_2) \dots V(t_n) \} \Phi(-\infty) \end{aligned} \quad (21.37)$$

and is identical with the previously obtained Equation (20.15).

CHAPTER V

THE S MATRIX

§ 22. Calculation of the S Matrix Elements.

1. The S Matrix.

Let us now consider the following general question. Assume that at the "initial" time $t = -\infty$ we are given a state of the electron-positron and electromagnetic fields; in other words, at $t = -\infty$ we are given the number of electrons, positrons, and photons in various individual states. We wish to find the state of the fields at $t = \infty$. This problem is most easily handled with the aid of the interaction representation, in which the state of the fields is described by the wave function $\Phi(t)$ which satisfies (20.6) and (21.32). We take the general solution (21.36) or (20.7), (20.15), and set $t = \infty$, obtaining

$$\Phi(+\infty) = P \left(e^{-i \int_{-\infty}^{\infty} V(t) dt} \right) \Phi(-\infty). \quad (21.1)$$

We see that the operator

$$S = P \left(e^{-i \int_{-\infty}^{\infty} V(t) dt} \right) \quad (22.2)$$

transforms the initial state of our system, which is given by $\Phi(-\infty)$, into the final state, described by $\Phi(+\infty)$. Thus, the problem reduces to determining S , called the scattering matrix.

The final state $\Phi(+\infty)$ can be considered a superposition of a set of mutually orthogonal states χ . If we are interested in the transition from the state Φ_0 to the state χ , we obtain the probability of such transition from $(\chi, \Phi(+\infty))$, which can be written, according to (22.1),

$$(\chi, \Phi(+\infty)) = (\chi, S \Phi(-\infty)). \quad (22.3)$$

Thus, the probabilities for various separate processes are determined by those elements of the S matrix which connect the appropriate initial and final states.

The S matrix can be expanded in a power series in the electron charge e . Clearly this expansion can be written¹⁾

$$S = \sum_{n=0}^{\infty} (-ie)^n \frac{1}{n!} \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n P(V^{(0)}(x_1) V^{(0)}(x_2) \dots V^{(0)}(x_n)), \quad (22.4)$$

where

$$V^{(0)}(x) = -j_{\mu}^{(0)}(x) A_{\mu}^{(0)}(x).$$

If we write the interaction energy $\mathcal{E} = -\int d^3x j_{\mu}^{(e)}(x) A_{\mu}^{(e)}(x)$ in the form of a sum

$$\mathcal{E} V^{(0)} = V^{(0)} + V^{(1)}, \quad (22.5)$$

where $V^{(0)}$ and $V^{(1)}$ are the interaction energies of the electron-positron field with the external fields (as well as photons) and the zero-point oscillations of the electromagnetic field, respectively, then it is easy to obtain a power series expansion in the external fields. This expansion can be written

$$S = S_0 + S_1 + \dots = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-i)^{m+n}}{m!n!} \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{m+n} \times \\ \times P(V^{(0)}(x_1) \dots V^{(0)}(x_m) V^{(1)}(x_{m+1}) \dots V^{(1)}(x_{m+n})). \quad (22.6)$$

¹⁾ We note that the various terms of (22.4) can be written

$$\int K_{\mu\nu} \dots (x_1, x_2, \dots) P(A_{\mu}(x_1) A_{\nu}(x_2) \dots) dx_1 dx_2 \dots,$$

where $K_{\mu\nu} \dots$ contains the chronological product of current density operators. Since $\int_{\mu}$ satisfies the continuity equation, we have

$$\frac{\partial K_{\mu\nu} \dots}{\partial x_{1\mu}} = \frac{\partial K_{\mu\nu} \dots}{\partial x_{2\mu}} = \dots = 0.$$

We used this expression in Section 16 for deriving Equation (16.21).

Here the zeroth order term S_0 , which does not contain $V^{(e)}(x)$, determines the scattering of electrons by the zero-point oscillations of the electromagnetic field. The first order term S_1

$$S_1 = -i \int_{-\infty}^{\infty} V_F(x) dx, \quad (22.7)$$

$$V_F(x) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n P(V^{(e)}(x) V^{(e)}(x_1) \dots V^{(e)}(x_n)),$$

determines the scattering due to the external field in the first Born approximation (for the external field) with the interaction of the electrons and positrons with the zero-point field oscillations taken into account. The succeeding terms of the series correspond to higher Born approximations with this same zero-point interaction taken into account.

We must bear in mind that unlike the interaction energy between the electron-positron field and the zero-point oscillations of the electromagnetic field, which can be treated as small perturbations, the interaction energy with the external field $V^{(e)}(x)$ can not always be considered a small perturbation. In particular, this cannot be done in determining the energy eigenvalues of the electron in an external field. In these cases one should use the wave functions of the electron in the external field to determine the current operator which enters into the expression for the S matrix.

2. Matrix Elements of the Field Operators.

Let us now establish the rules for calculating those S matrix elements which connect any two given states. We shall start with the general power series expansion in e as given by (22.4), in which the interaction energy $e V^{(e)}$ is not separated into $V^{(e)}$ and $V^{(i)}$ as described in the previous paragraph. The various terms of (22.4) are integrals whose integrands contain sums of products of the operators $\psi_e(x)$, $\bar{\psi}_e(x)$, $A_\mu(x)$.

It is easy to see that $\psi(x)$ is an electron annihilation and positron creation operator, $\bar{\psi}(x)$ is an electron creation and positron annihilation operator, and $A(x)$ is a photon emission and absorption operator, so that

$$\left. \begin{aligned} \psi(x) &= u(x) + \bar{v}(x), \\ \bar{\psi}(x) &= \bar{u}(x) + v(x), \\ A(x) &= a(x) + a^+(x), \end{aligned} \right\} \quad (22.8)$$

where u (or $\bar{u}$) is an electron annihilation (or creation) operator, v (or $\bar{v}$) is a positron annihilation (or creation) operator, and a (or a^+) is a photon absorption (or emission) operator.

To prove this assertion, let us consider those matrix elements of $\psi(x)$, $\bar{\psi}(x)$, and $A(x)$ which correspond to transitions from an initial state $\Phi(-\infty) = \Phi\{N_+, N_-, N\}$ to a final state $\Phi(+\infty) = \Phi\{N'_+, N'_-, N'\}$, where N_+ , N_- , N are the sets of electron, positron, and photon occupation numbers, respectively, in the initial state, and N'_+ , N'_- , N' are the sets for the final state. Equation (17.8) can be used to write the matrix elements of $\psi(x)$ and $\bar{\psi}(x)$ in the form

⁹ Since henceforth only the free-field operators enter the expressions, we shall not use the index zero as we did in Section 21.

$$\begin{aligned} \langle \Phi(+\infty), \psi(x) \Phi(-\infty) \rangle &= \frac{1}{\sqrt{V}} \left\{ \sum_p \sum_{r=1}^2 \bar{u}^r(p) e^{-ipx} \langle \Phi(+\infty), a_r(p) \Phi(-\infty) \rangle + \right. \\ &\quad \left. + \sum_p \sum_{r=1}^2 \bar{v}^r(p) e^{-ipx} \langle \Phi(+\infty), b_r^+(p) \Phi(-\infty) \rangle \right\}; \end{aligned} \quad (22.9)$$

$$\begin{aligned} \langle \Phi(+\infty), \bar{\psi}(x) \Phi(-\infty) \rangle &= \frac{1}{\sqrt{V}} \left\{ \sum_p \sum_{r=1}^2 \bar{u}^r(p) e^{-ipx} \langle \Phi(+\infty), a_r^+(p) \Phi(-\infty) \rangle + \right. \\ &\quad \left. + \sum_p \sum_{r=1}^2 \bar{v}^r(p) e^{-ipx} \langle \Phi(+\infty), b_r(p) \Phi(-\infty) \rangle \right\}. \end{aligned} \quad (22.10)$$

The matrix elements of $a_\pm(p)$, $a_\pm^+(p)$, $b_\pm(p)$, $b_\pm^+(p)$ are given by (17.13). If in the initial state there is only one electron with momentum p and polarization r , the only term different from zero in (22.9) will be

$$\langle \Phi(+\infty), a_r(p) \Phi(-\infty) \rangle = (a_r(p))_{01} = 1, \quad (22.11)$$

which means that in the final state the number of electrons with momentum p and polarization r is zero. If in the initial state there was no positron, but in the final state there is one positron with momentum p and polarization r , then the only term in (22.9) which does not vanish is

$$\langle \Phi(+\infty), b_r^+(p) \Phi(-\infty) \rangle = (b_r^+(p))_{10} = 1. \quad (22.12)$$

In other words $\psi(x)$ is an electron annihilation and positron creation operator, as stated above. Similarly, it can be shown that $\bar{\psi}(x)$ is a positron annihilation and electron creation operator.

It follows from (22.9), (22.11), and (22.12) that the matrix elements for annihilation of an electron with momentum p and creation of a positron with momentum p are, respectively,

$$\langle \Phi_{b_r^+}(p)(+\infty), \psi(x) \Phi_{a_r}(p)(-\infty) \rangle = \frac{1}{\sqrt{V}} \bar{u}^r(p) e^{ipx}, \quad (22.13)$$

$$(\Phi_{1_r^-(p)}(+\infty), \bar{\Phi}(x) \Phi_{0_r^-(p)}(-\infty)) = \frac{1}{\sqrt{V}} \bar{v}^r(p) e^{-ipx}. \quad (22.14)$$

These matrix elements are normalized Dirac plane waves with momenta $\mathbf{p}$ and $-\mathbf{p}$ corresponding to positive and negative energies.

Similarly, it can be shown that the matrix elements for creation of an electron with momentum $\mathbf{p}$ and annihilation of a positron with momentum $\mathbf{p}$ are, respectively

$$(\Phi_{1_r^+(p)}(+\infty), \bar{\Phi}(x) \Phi_{0_r^+(p)}(-\infty)) = \frac{1}{\sqrt{V}} \bar{u}^r(p) e^{-ipx}, \quad (22.15)$$

$$(\Phi_{0_r^-(p)}(+\infty), \bar{\Phi}(x) \Phi_{1_r^-(p)}(-\infty)) = \frac{1}{\sqrt{V}} \bar{v}^r(p) e^{ipx}. \quad (22.16)$$

Equations (22.15) and (22.16) are based on an expansion of ψ and $\bar{\psi}$ in plane waves. Such expansions, as has been noted in Section 17, can be used only when there is no external field or when the external field can be treated as a small perturbation. In general ψ should be expanded, as in (17.6) in electron eigenfunctions in the external field. In this case the annihilation and creation operators for the electron are given by

$$\left. \begin{aligned} (\Phi_{0_n^+}(+\infty), \bar{\Phi}(x) \Phi_{1_n^+}(-\infty)) &= \psi_n^{(+)}(x), \\ (\Phi_{1_n^+}(+\infty), \bar{\Phi}(x) \Phi_{0_n^+}(-\infty)) &= \bar{\psi}_n^{(+)}(x), \end{aligned} \right\} \quad (22.17)$$

where $\psi_n^{(++)}$ is the normalized wave function of the electron in the n -th state. Similar formulas are obtained for annihilation and creation of positrons.

Let us now consider the matrix elements of $A_\mu(x)$. Using the plane wave expansion (15.33), $(\Phi(+\infty), A_\mu(x) \Phi(-\infty))$ can be written¹⁾

$$\begin{aligned} (\Phi(+\infty), A_\mu(x) \Phi(-\infty)) &= \frac{1}{\sqrt{V}} \left\{ \sum_{k,\lambda} \frac{1}{\sqrt{2\omega}} e_{k\lambda} e^{ikx} (\Phi(+\infty), c_{k\lambda} \Phi(-\infty)) + \right. \\ &\quad \left. + \sum_{k,\lambda} \frac{1}{\sqrt{2\omega}} e_{k\lambda} e^{-ikx} (\Phi(+\infty), c_{k\lambda}^* \Phi(-\infty)) \right\}. \end{aligned} \quad (22.18)$$

The nonzero matrix elements of $c_{k\lambda}$ and $c_{k\lambda}^*$ are, according to (15.24),

¹⁾ [In equations (22.18)-(22.22) only the k 's appearing beneath the two summation signs should be boldface - editor's note.]

$$\left. \begin{aligned} (c_{k\lambda})_{N_{k\lambda}-1, N_{k\lambda}} &= \sqrt{N_{k\lambda}}, \\ (c_{k\lambda}^*)_{N_{k\lambda}+1, N_{k\lambda}} &= \sqrt{N_{k\lambda}+1} e \end{aligned} \right\} \quad (22.19)$$

Therefore, the matrix elements of $A_\mu(x)$, corresponding to absorption or emission, respectively, of a photon whose momentum is $\mathbf{k}$ and whose polarization is $\mathbf{e}_\mu$, are given by

$$(\bar{\Phi}_{0_k}(+\infty), A_\mu(x) \bar{\Phi}_{1_k}(-\infty)) = \frac{1}{\sqrt{2\omega V}} e_\mu e^{ikx}, \quad (22.20)$$

$$(\bar{\Phi}_{1_k}(+\infty), A_\mu(x) \bar{\Phi}_{0_k}(-\infty)) = \frac{1}{\sqrt{2\omega V}} e_\mu e^{-ikx}. \quad (22.21)$$

If, instead of a plane wave expansion for the electromagnetic field, we use a spherical wave expansion, then (22.20) and (22.21) can be replaced by

$$\left. \begin{aligned} (\bar{\Phi}_{0_k}(+\infty), A_\mu(x) \bar{\Phi}_{1_k}(-\infty)) &= (A_{JM})_\mu, \\ (\bar{\Phi}_{1_k}(+\infty), A_\mu(x) \bar{\Phi}_{0_k}(-\infty)) &= (A_{JM})_\mu^* \end{aligned} \right\} \quad (22.22)$$

where A_{JM} are the normalized momentum and parity eigenstates of the photon [see (5.19) and (5.21)].

Let us now return to the general expression (22.4) for the S matrix. Since the operators ψ , $\bar{\psi}$, A are sums of single-particle creation and annihilation operators, each term of (22.4) can be written as a sum of operators for the creation or annihilation of single electrons, positrons, and photons in various states. We must clarify the conditions under which this type of product has nonzero matrix elements corresponding to some definite process $i \rightarrow f$. If, for instance, there is one electron and no photon in state i , and there is a photon and an electron in state f , then obviously one of the annihilation operators annihilates the electron in state i , and two creation operators create the electron and photon in state f ; all the other operators can be separated into pairs, with the operators of each pair creating and annihilating the same particle.

The virtual processes of successive creation and annihilation of an individual particle make the calculation of the S matrix elements extremely complex. We shall, therefore, try to transform the S matrix to a form in which the virtual processes need not be considered. Clearly, the problem reduces to representing the S matrix as a sum of products of creation and annihilation operators, such that in each term the creation operators are on the left of the annihilation operators. In calculating the matrix elements of such products, the annihilation operators will annihilate only those particles which exist in the initial state, and the creation operators will create those which are in the final state. As for the virtual creation and annihilation processes, they will not enter explicitly into the considerations.

Products in which the creation operators are on the left of the annihilation operators shall be called ordered or normal products (compare Sections 16 and 18). It is clear that if we know the commutators and anticommutators of the field operators, we can always write the S matrix in the form of a sum of ordered products multiplied by some numerical coefficients. Ordering the field operators in the necessary way gives rise to additional terms containing commutators and anticommutators; these terms contain, though not in an explicit way, the virtual creation and annihilation processes.

3. Representation of the S Matrix as a Sum of Normal Products.

We shall now show that it is possible to represent the S matrix elements as sums of normal products of creation and annihilation operators by using algebraic techniques.¹⁾

First we shall give a general definition of the normal product of operators, which is a generalization of the definition of the ordered product given in Sections 16 and 18.

We shall denote the various terms of (22.8) which are either annihilation or creation operators by $U, V, W, \dots, Z$. If we are given a product of operators $UVW \dots Z$, we shall call $\delta_P XY \dots W$ the normal product corresponding to the given one, where $X, Y, \dots, W$ is the same set of operators as in the original product but ordered so that the creation operators are on the left of the annihilation operators, and the coefficient δ_P is $+1$ if the permutation of the electron-positron operators necessary to achieve this ordering is even, and is -1 if this permutation is odd. As for the various separate creation (or annihilation) operators, they may be permuted in any convenient way among themselves.²⁾ We shall denote the normal product by the symbol $\underline{N}$, so that

$$N(UV \dots Z) = \delta_P XY \dots W. \quad (22.23)$$

The normal product of operators $U \dots V \psi W \dots Z$, where $\psi = u + \bar{v}$ is a sum of a creation and an annihilation operator, is defined according to the distributive law

$$N(U \dots V \psi W \dots Z) = N(U \dots V u W \dots Z) + N(U \dots V \bar{v} W \dots Z).$$

The same law is valid if ψ is replaced by $\bar{\psi} = \bar{u} + v$ or $A = a + a^\dagger$.

Let us now consider the general expression (22.4) for the S matrix. Recalling [see (18.40)] that the current operator can be written in the form of an $\underline{N}$ -product

$$j_\mu(x) = ieN(\bar{\psi}(x)\gamma_\mu\psi(x))$$

and that, therefore,

$$eV^{(0)}(x) = -\int d^3x' A_\mu(x') j^\mu(x') = -ieN(\bar{\psi}(x)\hat{A}(x)\psi(x)), \quad (22.24)$$

¹⁾ G. Wick, Phys. Rev. 80, 268 (1950).

²⁾ This is due to the fact that the electron-positron operators anticommute, and the photon operators commute.

where

$$\hat{A}(x) = \gamma_\mu A_\mu(x),$$

we shall write the S matrix in the form

$$S = \sum_n S^{(n)},$$

$$S^{(n)} = \frac{(-e)^n}{n!} \int d^4x_1 \dots \int d^4x_n P \{ N(\bar{\psi}(x_1)\hat{A}(x_1)\psi(x_1)) \times$$

$$\times N(\bar{\psi}(x_2)\hat{A}(x_2)\psi(x_2)) \dots N(\bar{\psi}(x_n)\hat{A}(x_n)\psi(x_n)) \}. \quad (22.25)$$

The normal products $N(\bar{\psi}(x_1)\hat{A}(x_1)\psi(x_1))$ entering into this expression must be chronologically ordered.

It is convenient to use, instead of $\underline{P}(UV \dots Z)$, the so-called $\underline{T}$ -product, which differs from the $\underline{P}$ -product by the factor $\delta_P = \pm 1$, namely

$$T(UV \dots Z) = \delta_P XY \dots W, \quad (22.26)$$

where δ_P , as in (22.23), is determined by the permutation only of the electron-positron operators in (22.26); the operators $X, Y, \dots, W$ on the right side of (22.26) are chronologically ordered.³⁾

Since the electron-positron operators enter in pairs into (22.25), the operator $\underline{P}$ can be simply replaced by $\underline{T}$, so that

$$S^{(n)} = \frac{(-e)^n}{n!} \int d^4x_1 \dots \int d^4x_n T \{ N(\bar{\psi}(x_1)\hat{A}(x_1)\psi(x_1)) \dots N(\bar{\psi}(x_n)\hat{A}(x_n)\psi(x_n)) \}. \quad (22.27)$$

A $\underline{T}$ -product such as that in the integrand of (22.27), whose individual factors are $\underline{N}$ -products, will be called a mixed $\underline{T}$ -product. We shall now show that a mixed $\underline{T}$ -product can be written in the form of a sum of simple $\underline{N}$ -products of creation and annihilation operators. Let us consider the difference between the $\underline{T}$ - and $\underline{N}$ -product of two operators U and V (U and V can be $\psi, \bar{\psi}, A$ operators or their component parts $u, \bar{u}, v, \bar{v}, a, a^\dagger$). Let us denote this difference by $\underline{U} \underline{V}^C$, writing ⁴⁾

¹⁾ It can be shown that the $\underline{T}$ -product, unlike the $\underline{P}$ -product, is relativistically invariant.

²⁾ Instead of $\underline{C}$, we shall often use any other lower case Latin letter.

$$U^a V^a = T(UV) - N(UV), \quad (22.28)$$

which we call the contraction of the operators U and V .

In Sections 16 and 18 we determined one contraction of photon and electron-positron operators and saw that they are c -numbers.

If we know the contractions of the operators, it is simple to transform a T -product into a sum of N -products. This is done with the following two theorems.

I. If $U, V, \dots, X, Y, Z$ are $\psi, \bar{\psi}, A$ operators or their component parts $u, \bar{u}, v, \bar{v}, a, a^+$, then the T -product of these operators is the sum of their N -products with all possible contracted pairs; in other words,

$$T(UV \dots XYZ) = N(UV \dots XYZ) + N(U^a V^a W \dots XYZ) + N(U^a V^a W^b \dots XYZ) + \dots + N(U^a V^a W^b \dots X^c Y^c Z^c). \quad (22.29)$$

Here the different superscripts denote different contractions. The contraction of nonadjacent operators is defined in accordance with (22.23) as follows: if the contracted operators are photon operators, they are merely placed adjacent to each other; if they are electron-positron operators, they are placed adjacent, and the N -product must be multiplied by $\delta_{\pm}$ to indicate whether an even or odd permutation of the electron-positron operators is necessary.

For instance, if all the operators $U, V, \dots$ are electron-positron operators, then

$$N(U^a V W^b X^c Y^d Z) = -(U^a Y^d)(W^b X^c) N(VZ).$$

Here the contractions $U^a Y^d$ and $W^b X^c$ are c -numbers which we have taken out of the N -product. Similarly,

$$N(U^a V W^b X^c Y^d Z^e) = (U^a Y^d)(W^b Z^e) N(VX).$$

II. A mixed T -product, for instance $T(UV N(WXY) \dots Z)$, can be resolved into a sum of N -products similar to (22.29), except that the terms involving the contraction of operators within a given N -product should be dropped. In the example $T(UV N(WXY) \dots Z)$, we need not take into account the contractions $W^a X^a, W^a Y^a, X^a Y^a$.

Theorem I is proved as follows: note first that permutation simultaneously of the factors in the T -product on the left and the N -products on the right of (22.29) does not change this relation; we may, therefore, assume, without loss of generality, that the operators in (22.29) are chronologically ordered from right to left. If this is true, we shall say that the operators are T -ordered. Then the symbol T on the left side of (22.29) can be eliminated. Let us now order the operators on the left side of (22.29) so that all the creation operators are on the left of the annihilation operators. The operators are then called N -ordered. To do this let us take the furthest left N -unordered creation operator and interchange it successively with all annihilation operators on its left. We shall then obtain additional terms containing contractions between operators that have been interchanged, according to

$$UV = T(UV) = N(UV) + U^a V^a = \pm UV + U^a V^a.$$

Let us now perform this operation with the other unordered creation operators. We shall then have expressed the original T -product on the left side of (22.29) as a sum of N -products (we can clearly set the symbol N in front of each of the products). Although these N -products may enter both with positive and negative signs, if we reorder the factors within the N -products so that they are again T -ordered, then obviously all the N -products will enter with a positive sign. We then obtain an expression for the T -product in the form of a sum of N -products, which differs from (22.29) in that the right side will not contain all possible contractions between the factors, but only the contractions between pairs of N -unordered operators. Since the contraction between operators which are both N -ordered and T -ordered vanishes, we may add to the right side terms containing all possible contractions between pairs. This proves Theorem I.

Theorem II is proved similarly. In performing the proof, we need only bear in mind that it is unnecessary to interchange operators within a given N -product, since these operators are already N -ordered; therefore, these contractions do not enter into the expression.

In conclusion to this paragraph we present a summary of the formulas for the contractions of various operators:

$$\bar{\psi}_a^{\alpha}(x) \psi_b^{\beta}(y) = \frac{1}{2} S_{ab}^{\alpha\beta}(y-x); \quad (22.30)$$

$$\bar{\psi}_a^{\alpha}(y) \bar{\psi}_b^{\beta}(x) = -\frac{i}{2} S_{ab}^{\alpha\beta}(y-x); \quad (22.31)$$

$$A_a^{\alpha}(x) A_b^{\beta}(y) = \frac{1}{2} \delta_{ab} O^{\alpha\beta}(y-x); \quad (22.32)$$

$$\left. \begin{aligned} \bar{\psi}_a^{\alpha}(x) \bar{\psi}_b^{\beta}(y) &= 0, \\ \bar{\psi}_a^{\alpha}(x) \bar{\psi}_b^{\beta}(y) &= 0, \\ A_a^{\alpha}(x) \psi_b^{\beta}(y) &= 0. \end{aligned} \right\} \quad (22.33)$$

Equations (22.30) - (22.32) were obtained in Sections 16 and 18; the first two equations of (22.33) follow from (18.39); the third of equations (22.33) is self-evident.

Theorems I and II and these formulas simplify the problem of representing the integrand of expression (22.27) for the S matrix as a sum of normal products of creation and annihilation operators.

[All the ψ 's and $\bar{\psi}$'s in equations (22.30)-(22.33) should be boldfaced - editor's note.]

We note in conclusion that in the expression for the vacuum expectation value of products of the potential (16.25) or (22.32), both the longitudinal and scalar components enter similarly to the transverse ones. It is the former, as we shall see later, which cause the static (Coulomb) interaction of charges. Previously, before the development of covariant perturbation theory, the longitudinal and transverse degrees of freedom of the electromagnetic field were treated in essentially different ways. With the aid of a certain canonical transformation¹⁾ the longitudinal degrees of freedom were eliminated. Then terms corresponding to the Coulomb interaction appeared in the Hamiltonian, and the free electromagnetic field contained only transverse degrees of freedom. We shall not make use of this now obsolete method, since it lacks the advantage of relativistic invariance (which is extremely important for the elimination of divergences in the S matrix) and requires separate consideration of the static and retarded (due to emission and absorption of virtual photons) interactions. In the method we shall describe below for calculating the S matrix elements, the interaction of charges is characterized uniformly by the $\bar{\psi}\psi$.

§ 23. Graphic Representation of the Matrix Elements.

1. Graphic Representation of Normal Products.

In the preceding paragraph we have shown that the separate terms $S^{(n)}$ of the expansion of the S matrix in powers of the electron charge are integrals of mixed T-products of ψ , $\bar{\psi}$, A operators, and that these T-products can be resolved into normal products of the same operators. Each of the normal products into which the integrand of the $S^{(n)}$ matrix is resolved can be represented in the form of a diagram which is constructed in the following way.²⁾

The four-dimensional vectors $x_1, x_2, \dots, x_n$ over which the integration in the expression for $S^{(n)}$ is performed are represented by points on the diagram (these points shall be called vertices or corners of the diagram).

A contraction of photon operators A(x) and A(y) will be represented by a dotted line connecting the vertices x and y.

A contraction of operators $\bar{\psi}(x)$ and $\psi(y)$ will be represented by a solid line connecting the vertices x and y and directed from x to y.

The operator A(x) which is not contracted will be represented by a dotted line starting at x and leaving the diagram (going to "infinity").

Operators $\bar{\psi}(x)$ and $\psi(x)$ which are not contracted will be represented by solid lines from x out of the diagram; in the case of $\bar{\psi}(x)$, this line is directed from x to "infinity", whereas for $\psi(x)$, it is directed from "infinity" to x. Uncontracted operators shall be called free.

Since $\psi(x)$ is the electron annihilation and positron creation operator, a solid line directed from "infinity" to the vertex x is a graphic representation either of an electron which exists before the scattering process, or a positron created as a result of scattering.³⁾

Similarly, since $\bar{\psi}(x)$ is a positron annihilation operator and an electron emission operator, a solid line from a vertex x to "infinity" can represent both an electron created as a result of scattering, or a positron which existed before the scattering.

In addition, since A(x) is a photon emission and absorption operator, a dotted line connecting a vertex x with "infinity" can represent a photon either emitted or absorbed as a result of scattering.

A dotted line connecting a vertex x with "infinity" will also be used to represent the "external" electromagnetic field acting at the point x.

Lines connecting vertices of a diagram and representing contractions of operators can be interpreted in the following way. Since a contraction of operators contains a product of emission and absorption terms, lines connecting vertices can be associated with virtual particles⁴⁾ created or annihilated in the scattering process. An electron line is directed from its point of creation to its point of annihilation (and vice versa for a positron).

It is clear that at each vertex of a graph there appear two electron lines and one photon line.

Corresponding to the expansion of the S matrix in a series of powers of e, we shall say that a scattering process or an interaction process is an n-th order effect if the matrix element corresponding to this process is proportional to e^n . Obviously, all n-th order processes are described by the matrix $S^{(n)}$, the n-th term in the power series expansion of S. A diagram representing one of the normal products into which the integrand of the expression for $S^{(n)}$ is expanded contains n vertices. We shall call this an n-th order diagram.

A single diagram which represents some normal product of field operators can in general describe several different scattering processes (see, for instance, Diagram 3 of Fig. 10).

If normal products of operators can be described by diagrams which differ from each other only in the indices associated with the vertices, they are called equivalent. All these products clearly describe the same set of scattering processes (actually the same scattering process).

Let us consider, as an example, the matrix $S^{(3)}$, which describes third order scattering processes. Expanding the integrand into normal products according to Theorems I and II of Section 22, we obtain, among other terms,

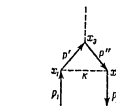


Fig. 8

$$-\frac{e^3}{3!} \int d^4x_1 \int d^4x_2 \int d^4x_3 N \{ \bar{\psi}^a(x_1) \hat{A}^b(x_1) \psi(x_1) \bar{\psi}(x_2) \times \\ \times \hat{A}^c(x_2) \psi^d(x_2) \bar{\psi}^e(x_2) \hat{A}(x_3) \psi^f(x_3) \}, \quad (23.1)$$

where the indices a, b, c, indicate the various contracted pairs.

The diagram corresponding to this term is shown in Fig. 8. Assuming that the external lines of the diagram represent an electron and an external electromagnetic field, we may say that this diagram represents the following process: an electron emits a virtual photon (at x_1) is then scattered by the external field A(x_2) (at x_2), and finally absorbs (at x_3) the virtual photon it emitted previously. To this process there correspond six equivalent terms in the decomposition of the integrand of $S^{(3)}$ into normal products. Dropping the symbol N and suppressing the arguments of the operators, these equivalent terms can be written

¹⁾ The momentum of a virtual electron, as opposed to a real one, does not satisfy the relation $E^2 - m^2 = 0$; the momentum of a virtual photon, as opposed to a real one, does not satisfy the relation $E^2 = 0$.

²⁾ E. Fermi, Revs. Mod. Phys. 4, 87 (1932).

³⁾ For more details see Chapter VII.

⁴⁾ This method was developed by Feynman (Phys. Rev. 76, 749, 769 (1940); a Russian translation can be found in the symposium Problems of Modern Physics Ser. 3, No. 11, pp. 25, 37).

⁵⁾ The possibility of such a representation of the motion of a positron (as an electron moving "backwards in time") was first discussed by G. Zisman [I. Exptl.-Theoret. Phys. (USSR) 10, 1063 (1940)].

$$\begin{aligned}
 1) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi); \\
 2) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi); \\
 3) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi); \\
 4) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi); \\
 5) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi); \\
 6) & (\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi)(\bar{\psi} \hat{A} \psi);
 \end{aligned}$$

diagram of Fig. 8 is six times the matrix element in (23.1).

Henceforth, we shall use $\underline{r}$ to designate the number of equivalent $\underline{N}$ -products in the integrand of the expression for $S^{(1)}$. Since the diagrams corresponding to these differ only in that the indices of the vertices are permuted, $\underline{r} = n!/g$, where g is the number of permutations of the indices which will not change the form of the $\underline{N}$ -product. For instance, in diagram 4 of Fig. 10, $g = 2$, and in diagram 6 of Fig. 23, $g = 4$.

2. Various Field Interaction Processes.

Let us go on to a construction of the diagrams and a calculation of the various interaction processes between the electron-positron and electromagnetic fields.

We shall start with first order effects. In this case, there is obviously only the one diagram shown in Fig. 9. It represents the scattering of an electron or positron in an external field, the emission or absorption of a photon by an electron (or positron), electron-positron pair creation or annihilation.

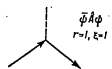


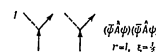
Fig. 9

On the right of each diagram is a symbolic representation of the integrand of the matrix element in $S^{(1)}$ (without the symbol $\underline{N}$ or the factor $e^{\pm 1/2}$). The lines connecting various of the factors $\bar{\psi}, \psi, A$ designate the contractions between the operators; the number $\underline{r}$ gives the number of $\underline{N}$ -products in the decomposition of $S^{(1)}$, and $\underline{\xi} = \frac{\underline{r}}{n!}$.

Figure 11 gives fifteen diagrams showing all possible third-order effects. Since the number of these effects is so great, we do not describe them, merely giving the expressions for the integrands of $S^{(3)}$, indicating the contracted pairs, and the values of $\underline{r}$ and $\underline{\xi} = \frac{\underline{r}}{3!}$.

where we have represented the contractions between electron operators by solid lines, and those between photon operators by dotted lines, as in the diagram. Diagrams corresponding to these terms will obviously differ only in the indices associated with the points $\underline{x}_i$.

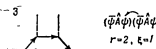
Noting that we can move the electron operators about in the $\underline{N}$ -product (multiplying, at the same time, by $\delta_{\underline{p}} = \pm 1$) and making use of (22.30) and (22.31), it is easily shown that all six terms are equal. Thus, the matrix element corresponding to the process represented by the



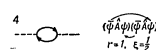
Two simultaneously occurring first order effects.



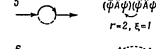
Scattering of an electron by an electron (or positron by a positron), or an electron by a positron.



Scattering of a photon by an electron; emission of two photons; two-fold electron (or positron) scattering in an external field; electron (or positron) bremsstrahlung; pair creation; two-photon pair annihilation.



Interaction between a photon and the electron-positron vacuum (photon "self-energy").



Interaction between an electron and the zero-point oscillations of the electromagnetic field (electron "self-energy").



Creation of a virtual pair and a virtual photon followed by annihilation of the pair and the photon (vacuum fluctuation).

Fig. 10

The diagrams under the number 10 differ in the direction of the arrows around the electron loop; according to Furry's theorem (see Section 24, below) these diagrams need not be considered, since the matrix element corresponding to them vanishes; for the same reason diagrams 15 can also be ignored.

The diagrams presented in Figs. 10 and 11 represent normal products of the field operators $A, \psi, \bar{\psi}$ in general form, and illustrate many processes. These operators represent the sums of creation and annihilation operators for particles in various states. Therefore, in considering any concrete physical process, the normal product which corresponds to it in the integrand of the expression for $S^{(n)}$ (that is, the normal product for the process being investigated) can in general be broken up into several terms each of which contains products of creation and annihilation operators for the particles taking part in the process being considered. These terms, which differ in the order in which the creation and annihilation operators of various particles appear, can also be represented by diagrams which are topologically equivalent and differ from each other only in the order in which the electron and photon lines occur in the diagram.

Let us consider, for instance, the emission of a k photon in an external field $A^{(e)}(x)$. This is a second order process, and the normal product corresponding to it in the integrand of the expression for $S^{(2)}$ can be written

$$N[\bar{\psi}(x_1) \hat{A}(x_1) \psi(x_1) \bar{\psi}(x_2) \hat{A}(x_2) \psi(x_2)]$$

In this expression $\hat{A}$ should be written

$$\hat{A}(x) = \hat{A}^{(e)}(x) + \hat{A}^{(k)}(x) + \hat{A}'(x),$$

where $\hat{A}^{(e)}$ is the external field, $\hat{A}^{(k)}$ is the emission and absorption operator of the k photon, and $\hat{A}'$ is the sum of similar operators referring to other photons. The normal product then breaks up into several terms, but the nonzero matrix elements for the process will clearly have only two terms, namely

$$N[\bar{\psi}(x_1)\hat{A}^{(e)}(x_1)\psi(x_2)\bar{\psi}(x_3)\hat{A}^{(k)}(x_3)\psi(x_4)] + N[\bar{\psi}(x_1)\hat{A}^{(e)}(x_1)\psi(x_2)\bar{\psi}(x_3)\hat{A}^{(k)}(x_3)\psi(x_4)].$$

Each of these terms can be represented graphically, using diagram 3 of Fig. 10, which represents the general normal product before the concrete form for $\hat{A}^{(k)}$ has been inserted into it. These diagrams are shown in

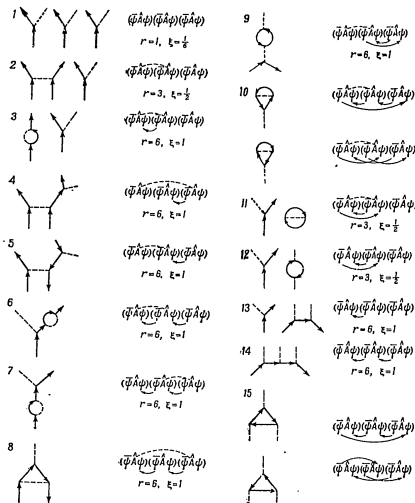


Fig. 11

Fig. 12, and differ from each other only in the order in which the photon lines corresponding to $\hat{A}^{(e)}$ and $\hat{A}^{(k)}$ appear.

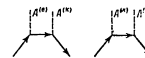


Fig. 13

If we were interested in a process in which three photons take part, then after the normal product was broken up into terms containing emission and absorption operators of the separate photons, we would obtain six terms for the nonzero matrix elements. These terms can be graphically represented, and the diagrams corresponding to them will differ only in the order of their photon lines.

Similarly, if several electrons or positrons take part in the process, the normal product can be represented as a sum of separate terms, each of which contains creation and annihilation operators for the electrons and positrons participating in the process, and which differ in the order of these operators. The various terms of the normal product can be represented by diagrams which are topologically equivalent and which differ only in the order of the lines corresponding to the various electrons and positrons.

In contradistinction to the case of several photons, in which the various graphs correspond to matrix elements all of which have the same sign, in the case of several electrons the various graphs correspond to matrix elements which may have different signs. This is related to the fact that the electron and positron annihilation and creation operators anticommute with each other, whereas those for the photons commute (we are assuming that the electrons are in different states).

This way of breaking up the normal product into separate terms containing annihilation and creation operators for particles in different states can be used in the general case when several photons and electrons participate in the process. Then the matrix element corresponding to the process under investigation is written in the form

$$S_{fi}^{(n)} = \sum M_{fi}^{(n)}, \quad (23.2)$$

where the individual terms of the sum differ from each other in the order of the creation and annihilation operators for the particles of interest, and may also differ in their sign. As was explained above, the diagrams corresponding to the individual terms of (23.2) are topologically equivalent, and differ only in the order of their electron and photon lines.

§ 24. The S Matrix in the Momentum Representation.

1. General Formula.

To determine the probabilities for various processes with the aid of the S matrix, as well as for a general examination of its properties, it is convenient to go to momentum space.

We shall start from Equation (23.2), which defines the matrix element $S_{fi}^{(n)}$ for an arbitrary process

$i \rightarrow f$ by means of a sum of terms $M_{fi}^{(n)}$ each of which contains matrix elements of annihilation and creation operators for the particles participating in the process under investigation. The $M_{fi}^{(n)}$ differ among themselves in the order in which these operators appear and are represented by diagrams which differ in the order of the electron and photon lines.

The matrix elements of the annihilation and creation operators are given by Equations (22.13)–(22.16), (22.20), and (22.21). We shall write these equations here again, assuming that the normalizing volume is $V = 1$. The matrix elements shall be denoted by the symbol of the corresponding quantity, with a superscript (+) or (−) depending on whether the particle is an electron or a positron.

The matrix elements of $\psi(\underline{x})$ and $\bar{\psi}(\underline{x})$ corresponding to annihilation and creation of an electron with momentum $\underline{p}$ are given by

$$\psi^{(+)}(\underline{x}) = u^r(\underline{p}) e^{i p \underline{x}}, \quad \bar{\psi}^{(+)}(\underline{x}) = \bar{u}^r(\underline{p}) e^{-i p \underline{x}}. \quad (24.1)$$

Similarly, the matrix elements of $\psi(\underline{x})$ and $\bar{\psi}(\underline{x})$ corresponding to creation and annihilation of a positron with momentum $\underline{p}$ are given by

$$\psi^{(-)}(\underline{x}) = v^r(\underline{p}) e^{-i p \underline{x}}, \quad \bar{\psi}^{(-)}(\underline{x}) = \bar{v}^r(\underline{p}) e^{i p \underline{x}}. \quad (24.2)$$

Here $u_{\alpha}^r(\underline{p})$ and $v_{\alpha}^r(\underline{p})$ are constant spinors satisfying the normalization conditions (17.9).

The matrix elements of $A_{\mu}(\underline{x})$ corresponding to absorption and emission of a photon with a four-momentum $\underline{k}$ and a polarization vector $\underline{e}$ are given, respectively, by

$$\begin{aligned} A_{\mu}(\underline{x}) &= \frac{1}{\sqrt{2\omega}} e_{\mu} e^{i k \underline{x}}, \\ A_{\mu}^*(\underline{x}) &= \frac{1}{\sqrt{2\omega}} e_{\mu} e^{-i k \underline{x}}. \end{aligned} \quad (24.3)$$

Let us establish the form of the $M_{\underline{1} \rightarrow \underline{f}}^{(\underline{n})}$ in momentum space. For this purpose we shall represent the contractions of operators, i.e., the functions $1/\hbar S^{\underline{F}}(\underline{x})$ and $1/\hbar D^{\underline{F}}(\underline{x})$, as well as the external potential $\underline{A}^{(\underline{e})}(\underline{x})$ in the form of Fourier integrals, and we shall insert these together with expressions (24.1)-(24.3) into Equation (22.25) for the matrix $S^{(\underline{n})}$; more exactly speaking, we shall insert these expressions into that one of the normal products, in the decomposition of the integrand of the expression for $S^{(\underline{n})}$, which corresponds to $M_{\underline{1} \rightarrow \underline{f}}^{(\underline{n})}$.

The contractions of the operators are given by the Fourier integrals

$$\begin{aligned} \frac{1}{2} S^{\underline{F}}(\underline{x}) &= \int \frac{1}{2} S^{\underline{F}}(\underline{p}) e^{i p \underline{x}} d^4 p, \\ \frac{1}{2} D^{\underline{F}}(\underline{x}) &= \int \frac{1}{2} D^{\underline{F}}(\underline{p}) e^{i p \underline{x}} d^4 p, \end{aligned} \quad (24.4)$$

where, according to (18.34) and (16.33),

$$\begin{aligned} \frac{1}{2} S^{\underline{F}}(\underline{p}) &= -\frac{i}{(2\pi)^4} \frac{(i \not{p} \not{p}_0 - m)_{\alpha\beta}}{p^2 + m^2}, \\ \frac{1}{2} D^{\underline{F}}(\underline{p}) &= -\frac{i}{(2\pi)^4} \frac{1}{p^2}, \end{aligned} \quad (24.5)$$

$(p^2 = p^2 - p_0^2).$

The Fourier expression for the external potential $\underline{A}^{(\underline{e})}_{\mu}(\underline{x})$, which we shall treat as a $\underline{c}$ -number, shall be written

$$\begin{aligned} A_{\mu}^{(\underline{e})}(\underline{x}) &= \frac{1}{(2\pi)^4} \int a_{\mu}(\underline{q}) e^{i q \underline{x}} d^4 q, \\ a_{\mu}(\underline{q}) &= \int A_{\mu}^{(\underline{e})}(\underline{x}) e^{-i q \underline{x}} d^4 x. \end{aligned} \quad (24.6)$$

Inserting (24.4), (24.6), and (24.1)-(24.3) into the normal product corresponding to $M_{\underline{1} \rightarrow \underline{f}}^{(\underline{n})}$, we shall first integrate over $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$. Collecting all factors $e^{i \underline{p} \underline{x}_1}$ with a given $\underline{x}_1$ (here $\underline{p}$ are the four-vector momenta of the free particles, as well as the variables of integration in the Fourier representations of $\underline{A}^{(\underline{e})}_{\mu}$, $\underline{S}^{\underline{F}}$ and $\underline{D}^{\underline{F}}$), we obtain $e^{i(\Sigma \underline{p}) \underline{x}_1}$, where the number of vectors in the sum $\Sigma \underline{p}$ is obviously three, i.e., the number of lines which meet at the vertex $\underline{x}_1$ of the diagram.

The integral of $e^{i(\Sigma \underline{p}) \underline{x}_1}$ is $(2\pi)^4 \delta(\Sigma \underline{p})$, where $\delta(\Sigma \underline{p})$ is the four-dimensional δ -function. Thus, in integrating over $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$, we obtain a product of n four-dimensional δ -functions.

We note that each line of the diagram corresponds to some four-dimensional vector $\underline{p}$. Since $\underline{S}^{\underline{F}}$ and $\underline{D}^{\underline{F}}$, which are related to internal lines of the diagram, depend on the differences between the coordinates at the ends of these lines, two of the δ -functions corresponding to some internal line contain the vector $\underline{p}$ associated with this line, but with opposite sign. This makes it possible to interpret the vector $\underline{p}$ corresponding to an internal line as the four-dimensional "momentum" of a virtual¹⁾ particle "emitted" at one end and "absorbed" at the other end of the internal line. The external lines of the diagram, that is those lines which leave the diagram, correspond to the four-dimensional momenta of real particles taking part in the process.

In order to obtain a final expression for $M_{\underline{1} \rightarrow \underline{f}}^{(\underline{n})}$, we must now integrate over the momenta $\underline{p}$. Denoting the momenta related to internal electron and photon lines of the diagram by $\underline{p}_i$ and $\underline{k}_i$, respectively, we obtain the following general expression in the form of a momentum-space integral for $M_{\underline{1} \rightarrow \underline{f}}^{(\underline{n})}$:

¹⁾ We recall that for virtual particles there exists no definite relation between the time and space components of $\underline{p}$ (see the note on p. 241).



Fig. 13

a contraction of the type $\bar{\psi}^c \psi^c$; since there are three contractions, the sum of the matrix elements for both diagrams vanishes.

Furry's theorem follows from the conservation of charge parity (see Section 19, paragraph 3). Indeed, a closed electron loop connected to n photon lines corresponds to a neutral system of photons (real and virtual), whose total number, both initial and final, is n . Since $\Lambda = -1$ for a single photon, odd n would give different parities in the initial and final states.

3. Summary of the Rules.

Let us now summarize. We see that Equation (24.7), which determines $M_{i \rightarrow f}^{(n)}$, can be written down immediately for each of the diagrams if we bear in mind the following rules:

1) Each external electron line corresponds to one of the spinors $u, \bar{u}, v, \bar{v}$, where u and $\bar{u}$ correspond to annihilation and creation of an electron, and v and $\bar{v}$ correspond to annihilation and creation of a positron. The four-momentum of the positron is the negative of that given by the defining spinor v (v corresponds to a plane wave with negative frequency).

2) Each external photon line which corresponds to a photon is associated with an operator $\frac{\hat{a}}{\sqrt{2\omega}}$. Each external photon line which corresponds to the external electromagnetic field is associated with an operator $\frac{1}{(2\pi)^4} \hat{a}(q)$.

3) Each internal electron line is associated with an operator $\frac{i\hat{p} - m}{p^2 + m^2}$.

4) Each internal photon line is associated with a factor $\frac{1}{k^2}$, and its ends are associated with γ_ν matrices.

5) Each vertex of the graph is associated with $\delta(\Sigma p)$, containing the momentum of all the lines which meet at this vertex (the momenta at opposite ends of internal lines should be taken with opposite signs).

6) All the operators (acting on the spinor indices) are ordered from right to left as they are encountered when moving in the direction of an electron line.

7) If a diagram contains a closed electron loop, the expression for $M_{i \rightarrow f}^{(n)}$ contains the trace of the product of the matrices $\frac{i\hat{p} - m}{p^2 + m^2}$ and γ_ν referring to the separate lines of the loop and to its vertices.

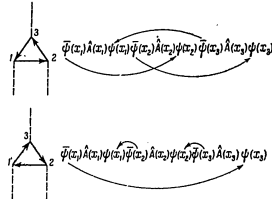


Fig. 14

The matrix elements of diagrams containing one or more closed electron loops with odd numbers of electron lines vanish identically.

8) Integration is performed over the four-momenta of the internal lines representing virtual particles and over the variables q of the external potentials; summation is taken over the four-dimensional polarizations of the virtual photons.

9) The numerical coefficient multiplying the general expression for $M_{i \rightarrow f}^{(n)}$ is $\pm \epsilon \frac{n}{n!} (-1)^F (2\pi)^4 \delta^{(n-F)}$. Here $\epsilon = \frac{1}{n!}$, F is the total number of internal lines, and the sign is chosen to correspond with the rule for permuting the factors in an N -product and Equations (22.30) and (22.31) for contracted electron-positron operators.¹⁾

As an illustration of these rules let us determine the matrix element for scattering of an electron in an external field with emission and absorption of a virtual photon (see Section 42). This process is described by the diagram shown in Fig. 13. In the general expression (23.2) for $S_{i \rightarrow f}^{(n)}$, only one term $M_{i \rightarrow f}^{(n)}$ is relevant in this case, and according to (24.7) it can be written

$$S_{i \rightarrow f}^{(1)} = (-i)^3 e^2 \bar{u}_2 \gamma_\nu u_1 \int \frac{d^4 k}{(2\pi)^4} \frac{1}{(p_1 - k)^2 + m^2} \frac{1}{(p_2 - k)^2 + m^2} \gamma_\nu \frac{1}{k^2} u_1, \quad (24.8)$$

where u_1 and u_2 are spinor amplitudes with four-momenta p_1 and p_2 of the electron before and after scattering, and $q = p_2 - p_1$. The three four-dimensional δ -functions which, according to (24.7), enter the expression for $M_{i \rightarrow f}^{(1)}$ have been eliminated by integration over q and over the momenta of the internal electron lines, replacing q by $p_2 - p_1$, and the momenta of the internal electron lines by $p_1 - k$ and $p_2 - k$. For this reason, only integration over the four-momentum k of the virtual photon appears in (24.8). This equation involves also summation over the polarizations of the virtual photon, i.e., over the index ν .

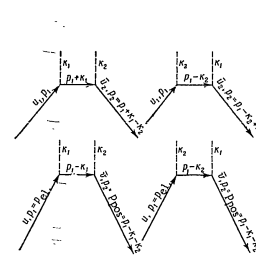


Fig. 15

(the corresponding spinor amplitude in Fig. 15a is denoted by $\bar{u}_2$), and in case b it corresponds to the annihilation of the positron with four-momentum $p_{\text{pos}} = -p_2$ (the spinor amplitude corresponding to the momentum p_2 is denoted by $\bar{v}$ in Fig. 15b).

¹⁾ The rules formulated above were established by Feynman. See the references given on p. 24.

According to Equation (24.7) and the above rules, the matrix elements for the Compton effect and for two-photon pair annihilation can be written

$$S_{\text{scat}}^{(2)} = M_{\text{scat}, I}^{(2)} + M_{\text{scat}, II}^{(2)} = -ie^2 (2\pi)^4 \bar{u}_s \left(\frac{\epsilon_2}{\sqrt{2\omega_2}} \frac{i(\hat{p}_1 + \hat{k}_1) - m}{(p_1 + k_1)^2 + m^2} \frac{\epsilon_1}{\sqrt{2\omega_1}} + \right. \\ \left. + \frac{\epsilon_1}{\sqrt{2\omega_1}} \frac{i(\hat{p}_1 - \hat{k}_2) - m}{(p_1 - k_2)^2 + m^2} \frac{\epsilon_2}{\sqrt{2\omega_2}} \right) u_1 \delta(p_2 - p_1 + k_2 - k_1), \quad (24.9)$$

$$S_{\text{pair ann.}}^{(2)} = M_{\text{ann.}, I}^{(2)} + M_{\text{ann.}, II}^{(2)} = -ie^2 (2\pi)^4 \bar{v} \left(\frac{\epsilon_2}{\sqrt{2\omega_2}} \frac{i(\hat{p}_1 - \hat{k}_2) - m}{(p_1 - k_2)^2 + m^2} \frac{\epsilon_1}{\sqrt{2\omega_1}} + \right. \\ \left. + \frac{\epsilon_1}{\sqrt{2\omega_1}} \frac{i(\hat{p}_1 - \hat{k}_2) - m}{(p_1 - k_2)^2 + m^2} \frac{\epsilon_2}{\sqrt{2\omega_2}} \right) u \delta(p_2 - p_1 + k_1 + k_2), \quad (24.10)$$

where the indices I and II denote the separate diagrams of Fig. 15a and b, which differ in the order of the photon operators; ϵ_1 and ϵ_2 are the polarizations of the photons with momenta k_1 and k_2 . One of the two δ -functions which, according to (24.7), enter the expression for $M_{I \rightarrow f}^{(n)}$, has been eliminated by integration over the momentum of the internal electron line and by replacing it by $p_1 + k_1$ and $p_1 - k_2$ in diagram a, and by $p_1 - k_1$ and $p_1 + k_2$ in diagram b.⁹⁾

§ 25. Analysis of the Singularities of the S Matrix.

1. General Properties of the Diagrams.

The momentum representation of the S matrix makes it possible to undertake a general analysis of the singularities of its matrix elements. Before going on to a consideration of this problem, we shall make some preliminary general remarks on the structure of the diagrams representing the matrix elements.

If a diagram contains several disconnected parts, it obviously represents several unrelated scattering processes. The matrix element corresponding to the diagram then breaks up into a series of individual factors, each of which is the matrix element for an individual scattering process. Clearly it is sufficient to analyze those matrix elements and diagrams which do not break up into separate disconnected parts.

In many cases a diagram may contain one or more parts which are connected to the rest of the diagram only by two similar (electron or photon) lines. Such parts will be called self-energy parts.⁹⁾

⁹⁾ Sections 25-27 which follow are devoted to an investigation of the singularities of the S matrix and to a description of methods for removing divergences. These methods will be necessary in studying radiative correction in Chapter VIII. Therefore, the reader may skip Sections 25-27, and after studying Section 28 may go on to Chapter VI and VII which discuss concrete phenomena in the first approximation.

⁹⁾ This terminology is related to the fact that the self-energy parts describe both the interaction of the electron with zero-point oscillations of the electromagnetic field, and the interaction of the photon with the zero-point oscillations of the electron-positron field; see below.

A part of the diagram connected to the rest only by two electron lines will be called an electron self-energy part.

A part of the diagram connected to the rest only by two photon lines will be called a photon self-energy part.

A part of the diagram connected to the rest only by one photon and two electron lines will be called a vertex part.

Such parts are indicated schematically in Fig. 16, where they are denoted by W_e , W_p , V , respectively (the squares A, B, and C denote the rest of the diagrams).

The simplest W_e , W_p , and V structures are shown in Fig. 17.

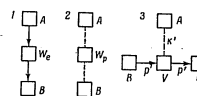


Fig. 16

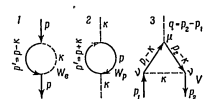


Fig. 17

We shall differentiate between two classes of internal (self-energy, vertex, etc.) parts, namely reducible and irreducible.

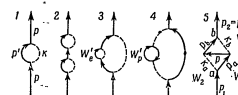


Fig. 18

An irreducible part is one which cannot be separated into parts connected by only one line and which does not contain self-energy or vertex parts.

If this is not the case, we call a self-energy, vertex, or other internal part of a diagram reducible.

Figure 18 shows second-order and fourth-order electron self-energy parts.

Diagram 1 of this figure is irreducible, and the rest are reducible. Of these, diagram 2 can be divided into two electron self-energy parts which are connected by a single electron line; diagram 3 contains an electron self-energy part W_e ; diagram 4 contains a photon self-energy part W_p .

part W_e ; diagram 5 contains a vertex part V_a at a , and another vertex part V_b at b (to V_a along the internal photon line k_a and the internal electron lines p_a and p_b , while the photon line k_b and the electron lines p_b and p_1 are external with respect to V_a ; similarly, V_b contains the internal photon line k_b and the internal electron lines p_b and p_1 , while the photon line k_a and the electron lines p_a and p_2 are external with respect to V_b).

It can be shown that diagram 1 is the only irreducible electron self-energy diagram.

Figure 19 shows second and fourth order photon self-energy parts.

Diagram 1 is irreducible, and the rest are reducible. Diagram 2 can be separated into two photon self-energy parts connected by a single photon line; diagram 3 contains an electron self-energy part W_e ; diagram 4 contains a vertex part V_a at a and a vertex part V_b at b (to V_a along the internal photon line k and the internal

electron lines p'_a and p''_a , while the photon line k_1 and the electron lines p'_b and p''_b are external to V_b ; similarly, to V_b belong the internal photon line k and the internal electron lines p'_b and p''_b , while the photon line k_2 and the electron lines p'_a and p''_a are external to V_a .

Figure 18, diagram 5 and Fig. 19, diagram 4 show that if an electron or photon self-energy part has a vertex part at a , then there is also a vertex part at b .

It can be shown that diagram 1 of Fig. 19 is the only irreducible photon self-energy diagram.

Figure 20 shows third and fifth order vertex parts; of these, diagrams 1 and 2 are irreducible, and the rest are reducible.

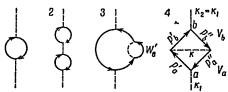


Fig. 19

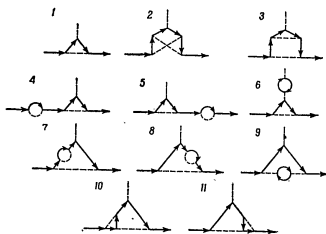


Fig. 20

2. Equivalent Skeleton Diagrams.

The reason for giving special consideration to the self-energy and vertex parts is related to the fact that in the general expression (24.7) for $\frac{M^{(n)}}{i} \rightarrow i$ these parts contribute separate factors which do not depend on the structure of the rest of the diagram (that is, on the structure of A , B , and C of Fig. 16); therefore, if they have been previously calculated, they can be used to determine the matrix elements of various processes.

In doing this, the diagrams may be replaced by certain "equivalent" skeleton diagrams which do not contain self-energy and vertex parts; the expressions for the contraction of operators and other quantities related to the lines and vertices of the diagram from which the self-energy and vertex parts have been eliminated are then modified.

Skeleton diagrams are shown in Fig. 21.

The electron line connecting A and B from which the electron self-energy part has been eliminated, now corresponds not to $\frac{1}{2} S^F$, but to some other operator $\frac{1}{2} \delta S^F$ which depends on the momentum p of the electron line and on the actual form of the electron self-energy part W_e . If W_e is an irreducible electron self-energy part, such as is indicated in Fig. 17, diagram 1, then according to (24.7) $\frac{1}{2} \delta S^F$ can be written

$$\frac{1}{2} \delta S^F(p) = \frac{e^2}{(2\pi)^3} \frac{i\hat{p} - m}{p^2 + m^2} \int \gamma, \frac{i(\hat{p} - \hat{k}) - m}{(p - k)^2 + m^2} \gamma, \frac{d^4k}{k^2} \frac{i\hat{k} - m}{k^2 + m^2}. \quad (25.1)$$

We see [see (24.5)] that

$$\frac{1}{2} \delta S^F(p) = \frac{1}{2} S^F(p) \Sigma(W_e, p) \frac{1}{2} S^F(p), \quad (25.2)$$

where

$$\Sigma(W_e, p) = e^2 \Sigma^{(0)}(W_e, p) = -e^2 \int \gamma, \frac{i(\hat{p} - \hat{k}) - m}{(p - k)^2 + m^2} \gamma, \frac{d^4k}{k^2}. \quad (25.3)$$

This formula is valid for the irreducible electron self-energy part, but $\frac{1}{2} \delta S^F$ can always be written in the form of (25.2), where $\Sigma(W_e, p)$ is some expression, more complex than (25.3), depending on the actual structure of W_e .

We note that the integral in (25.3) diverges linearly for large k^0 .

If part A of the diagram of Fig. 16 is absent, that is if the electron line connecting W_e with A describes a free electron or positron, then the electron line from which the self-energy part has been eliminated will correspond not to the spinor $u(p)$, but to some modified spinor $\delta u(p)$, which can be written

$$\delta u(p) = \frac{1}{2} S^F(p) \Sigma(W_e, p) u(p), \quad (25.4)$$

where $\Sigma(W_e, p)$ is the same operator as that in (25.2).

Similarly, the spinor $\bar{u}(p)$ must be replaced by

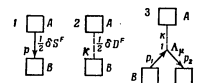


Fig. 21

⁹ We define the degree of divergence of an integral by the difference in the degrees of the numerator and denominator of the integrand.

$$\bar{u}(p) = \bar{u}(p) \Sigma(W_p, p) \frac{1}{2} S^F(p) \quad (25.4')$$

in a skeleton diagram. Similar expressions are obtained for the spinors $\bar{v}$ and $\bar{v}$.

Photon lines connecting A and B correspond, in skeleton diagrams, not to $\frac{1}{2} D^F$, but to some other function $\frac{1}{2} \delta D^F$ which depends on the momentum k of the photon line and on the actual form of the photon self-energy part $\underline{W}_P$.

If $\underline{W}_P$ is an irreducible photon self-energy part, such as is shown in Fig. 17, diagram 2, then according to (24.7) $\frac{1}{2} \delta D^F$ can be written

$$\frac{1}{2} \delta D^F(k) = - \frac{e^4}{(2\pi)^4} \frac{1}{k^2} \int \text{Spur} \left\{ \gamma_1 \frac{i\hat{p} - m}{p^2 + m^2} \gamma_1 \frac{i(\hat{p} + k) - m}{(p+k)^2 + m^2} \right\} d^4 p \frac{1}{k^2}. \quad (25.5)$$

We see [see (24.5)] that

$$\frac{1}{2} \delta D^F(k) = \frac{1}{2} D^F(k) \Pi(W_p, k) \frac{1}{2} D^F(k), \quad (25.6)$$

where

$$\Pi(W_p, k) = e^4 \Pi^{(0)}(W_p, k) = e^4 \int \text{Spur} \left\{ \gamma_1 \frac{i\hat{p} - m}{p^2 + m^2} \gamma_1 \frac{i(\hat{p} + k) - m}{(p+k)^2 + m^2} \right\} d^4 p. \quad (25.7)$$

This equation is valid for an irreducible photon self-energy part, but $\frac{1}{2} \delta D^F$ can always be written in the form of (25.6) where $\Pi(W_p, k)$ is given by an expression more complicated than (25.7), depending on the actual structure of $\underline{W}_P$.

We note that the integral in (25.7) diverges quadratically for large p .

If part A of the diagram is absent, that is if the photon line connecting $\underline{W}_P$ with $\underline{A}$ represents a free electron or an external electromagnetic field, then the photon line from which the self-energy part has been eliminated will correspond not to the operator $\hat{A}(k)$, but to some modified operator $\delta \hat{A}(k)$ given by

$$\delta \hat{A}(k) = \hat{A}(k) \Pi(W_p, k) \frac{1}{2} D^F(k), \quad (25.8)$$

where Π is the same expression as (25.6).

Finally, if the diagram contains a vertex part, then to point 1 of the skeleton graph (see Fig. 21, diagram 3) there corresponds not a γ_μ matrix, but some other operator Λ_μ depending on the form of the vertex part $\underline{V}$ and on the momenta p_1 and p_2 of the electron lines at point 1.

In the case of an irreducible third order vertex part, which is shown in Fig. 17, diagram 3, according to (24.7) p_1 and p_2 can be written in the form

$$\Lambda_\mu(V, p_1, p_2) = e^3 \Lambda_\mu^{(0)}(V, p_1, p_2) = \frac{ie^3}{(2\pi)^4} \int \gamma_1 \frac{i(\hat{p}_1 - k) - m}{(p_1 - k)^2 + m^2} \gamma_\mu \frac{i(\hat{p}_2 - k) - m}{(p_2 - k)^2 + m^2} \gamma_1 \frac{d^4 k}{k^2}. \quad (25.9)$$

We see that Λ_μ is given by an integral which diverges logarithmically for large k .

Thus, if a diagram contains self-energy and vertex parts, its matrix elements can be obtained from the matrix element of a skeleton diagram with the following substitutions:

$$\left. \begin{aligned} \frac{1}{2} S^F &\rightarrow \frac{1}{2} \delta S^F, & \frac{1}{2} D^F &\rightarrow \frac{1}{2} \delta D^F, \\ u, \bar{v} &\rightarrow \delta u, \delta \bar{v}, & \underline{u}, \bar{\underline{v}} &\rightarrow \delta \underline{u}, \delta \bar{\underline{v}}, \\ A_\mu &\rightarrow \delta A_\mu, & \gamma_\mu &\rightarrow \delta \gamma_\mu. \end{aligned} \right\} \quad (25.10)$$

Of course the analysis of internal component parts of diagrams need not necessarily be restricted to the consideration of only self-energy and vertex parts; we could investigate also parts of a diagram which are connected with the rest by a larger number of lines than in the above cases. It is then necessary, however, to bear in mind that if a diagram contains a part connected with the rest of the diagram by only an odd number of photon lines, the matrix element corresponding to it vanishes. This follows directly from Furry's theorem, since this part of the diagram contains at least one closed electron loop with an odd number of electron lines.

In analyzing the singularities of the S matrix, in addition to the self-energy and vertex parts, parts of the diagram connected to the rest by only four photon lines are also of importance. The simplest form of such a part is shown in Fig. 22. We shall call this a photon-photon scattering part, since if parts A, B, C, and D are absent, then the diagram corresponds to just this effect.

The self-energy and vertex parts lead to divergent matrix elements, since the operators δS^F , Λ_μ , δA_μ , as well as the spinors δu , $\delta \bar{u}$ and the function δD^F are given by integrals which diverge for high momenta of the virtual particles.⁵⁾

The rest of our problem consists of eliminating the divergences due to integration over infinite virtual particle momenta, and of obtaining finite quantities which have physical meaning from these divergent matrix elements.

We note that in many problems divergence arises also in the low-momentum region. This kind of divergence

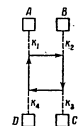


Fig. 22

⁵⁾ Let us recall that virtual particles are related to internal lines of the graphs, and that the time components of their four-momenta are not related to their space components.

is called the *infra-red catastrophe* and is due to the inapplicability of perturbation theory; it can be relatively simply eliminated (see, in this connection, Section 32). Therefore, we shall consider only the high-momentum virtual particle divergences.

3. Possible Types of Divergences Related to Irreducible Diagrams.

Before going on to a solution of this problem, we shall discuss the character of the divergences which arise. We note that in sufficiently complex self-energy and vertex parts, the matrix element may contain divergent factors which can, in general, lead to infinities of arbitrarily high orders. Whereas, for instance, the self-energy part of Fig. 18, diagram 1 leads to an infinity which we shall agree to call a first-order infinity, the self-energy part of Fig. 18, diagram 2 clearly leads to one of second order.

It follows from this that we must first consider divergences arising from irreducible internal parts of a diagram (that is, from irreducible self-energy and vertex parts). We shall show that the irreducible internal parts lead to just four types of divergent matrix elements.⁹⁾

Let n be the number of vertices of an irreducible internal part of a diagram, F be the number of its internal lines, and N be the number of lines external to this part and connecting it with the rest of the diagram (that is, with A , B , C of Fig. 16). Let N_e be the number of such external electron lines, and N_p the number of such external photon lines, so that $N = N_e + N_p$. Let us consider that part of the total matrix element which is related to the dynamic variables belonging to the internal parts of the diagram. This part of the matrix element is clearly an integral over $4F$ variables, namely the momenta of the internal lines of the irreducible internal part of the diagram. The momenta of the N external lines will enter this integral as constant parameters.

Since the integrand contains n four-dimensional δ -functions, not all of the $4F$ variables of integration will be independent. Clearly we can eliminate only $n-1$ of these δ -functions by integrating over $4(n-1)$ variables; one of the δ -functions, which expresses the conservation of the four-momentum in this part of the diagram, must remain. The conservation law expresses the relation between the momenta of the lines which connect the self-energy parts W_e , W_p and the vertex part V with the rest of the diagram. We may thus assert that the number of independent variables of integration will be $4(F-n+1)$.

The integrand is obviously a rational function, and the degree of the numerator, according to (24.7), will be F_e , and that of the denominator will be $2F$ (here F_e is the number of internal electron lines).

Since our diagram is irreducible, we may assert that the integrand does not break up into separate factors containing variables which are not related to each other. Thus, the convergence of the integral in the high-momentum region is determined by the difference in the degree of the numerator and that of the denominator, which can be written

$$K = 2F - F_e - 4(F - n + 1).$$

If $K \approx 1$, the integral converges; it diverges logarithmically when $K = 0$, linearly when $K = -1$, quadratically when $K = -2$, etc.

Since each corner of the graph contains two electron lines and one photon line, then clearly

$$2F_e + N_e = 2n, \quad 2F_p + N_p = n,$$

whence it follows that

$$K = \frac{3}{2}N_e + N_p - 4. \quad (25.11)$$

Equation (25.11) makes it possible to enumerate all possible cases of divergent irreducible internal parts of the diagrams. There exist, clearly, seven types of divergences connected with the following values for N_e and N_p :

- 1) $N_e = 2, N_p = 0, K = -1$, linear divergence;
- 2) $N_e = 2, N_p = 1, K = 0$, logarithmic divergence;
- 3) $N_e = 0, N_p = 1, K = -3$, cubic divergence;
- 4) $N_e = 0, N_p = 2, K = -2$, quadratic divergence;
- 5) $N_e = 0, N_p = 3, K = -1$, linear divergence;
- 6) $N_e = 0, N_p = 4, K = 0$, logarithmic divergence;
- 7) $N_e = 0, N_p = 0, K = -4$, fourth-order divergence.

Figure 23 shows the simplest irreducible diagrams corresponding to these divergences.

We recall that there exists only one irreducible electron self-energy diagram and one irreducible photon self-energy diagram; the number of irreducible vertex parts and photon-photon scattering parts is unlimited. Figure 23 shows the simplest types of these diagrams.

Furry's Theorem may be used to exclude cases 3 and 5 immediately, since the internal parts of these diagrams are connected to the rest by an odd number of photon lines, so that the corresponding matrix elements vanish.

Furthermore, case 7 clearly has no physical meaning, since in this case the internal part of the diagram is not connected with the rest of it. Diagram 7 obviously describes a vacuum-vacuum transition in the absence of external fields, and the matrix element corresponding to it gives a second order correction to the amplitude of this transition. On the other hand, it is clear that the probability for vacuum-vacuum transition in the absence of an external field is unity, and therefore, all corrections to it should be considered zero.

Thus, we see that there actually exist only four types of divergences related to irreducible electron and photon self-energy parts, irreducible vertex parts, and irreducible photon-photon scattering parts. We shall go on to a consideration of these divergences.

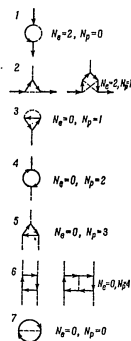


Fig. 23

⁹⁾ This result is due to Dyson, Phys. Rev. 75, 1738 (1949) (a Russian translation of this article is available in the symposium, Problems of Modern Physics, No. 11, p. 64).

§ 26. Removal of the Divergences From the S Matrix.

1. Removal of the Divergence Due to Photon-Photon Scattering.

Let us now ask the following question: Is it possible to remove a divergence from the S matrix and to obtain a finite expression with physical meaning from a divergent matrix element?

Let us first consider the divergence due to photon-photon scattering. The simplest diagram which describes this process is shown in Fig. 22. We shall assume that parts A, B, C, and D of the diagram are absent, and shall use the symbols $\Delta_\mu(k_1)$, $\Delta_\nu(k_2)$, $\Delta_\lambda(k_3)$, $\Delta_\sigma(k_4)$ to denote the four-dimensional potentials related to the photons with momenta k_1 , k_2 , k_3 , k_4 . The matrix element corresponding to this diagram, according to the general rules of Section 24, can be written

$$S_{\mu\nu\lambda\sigma}^{(0)} = A_\mu(k_1) A_\nu(k_2) A_\lambda(k_3) A_\sigma(k_4) I_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4) \delta(k_1 + k_2 + k_3 + k_4), \quad (26.1)$$

where $I_{\mu\nu\lambda\sigma}$ is an integral, over the momentum p of the virtual electron, of a rational function $Q_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4, p)$ of p and k_i .⁴⁾

$$I_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4) = \int Q_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4, p) d^4p.$$

(In the case of more complex diagrams, such as diagram 6 on the right in Fig. 23, $Q_{\mu\nu\lambda\sigma}$ is a function not only of k_i and p , but also of the momenta of the virtual photons, and the integration in the expression for $I_{\mu\nu\lambda\sigma}$ is taken also over the momenta of these photons.)

In the previous section we saw that $I_{\mu\nu\lambda\sigma}$ diverges logarithmically for large p . We shall now show how to eliminate this divergence. For this purpose we note that the part of $I_{\mu\nu\lambda\sigma}$ corresponding to large values of $|p|$ is independent of the k_i , since p and k_i appear as a sum, and $|p| \gg |k_i|$. Therefore, the corresponding part of $S_{\mu\nu\lambda\sigma}^{(0)}$ contains only the four-dimensional photon polarizations. This means that the expression for $I_{\mu\nu\lambda\sigma}$ for large $|p|$ depends only on the potentials, and not on their derivatives; thus, it is not gauge invariant. It then follows that the part of $S_{\mu\nu\lambda\sigma}^{(0)}$ which corresponds to the region of large $|p|$, which we shall denote by $S_{\mu\nu\lambda\sigma}^{(4)}$, can be ignored. This part should be separated out in a way such that the remaining part of $S_{\mu\nu\lambda\sigma}^{(0)}$ be finite and gauge invariant. Let us make use of the fact that when $k_1 = k_2 = k_3 = k_4 = 0$, then $S_{\mu\nu\lambda\sigma}^{(4)} = S_{\mu\nu\lambda\sigma}^{(0)}$. Thus, in order to obtain a finite and gauge invariant result from $I_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4)$, we must subtract the value of this integral when $k_1 = k_2 = k_3 = k_4 = 0$. We shall show that the quantity thus obtained, namely

$$I_{\mu\nu\lambda\sigma R}(k_1, k_2, k_3, k_4) = \int (Q_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4, p) - Q_{\mu\nu\lambda\sigma}(0, 0, 0, 0, p)) d^4p.$$

⁴⁾ An explicit expression for this function can be found in Section 47.

is finite. To do this, we note that

$$Q_{\mu\nu\lambda\sigma R}(k_1, k_2, k_3, k_4, p) = Q_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4, p) - Q_{\mu\nu\lambda\sigma}(0, 0, 0, 0, p)$$

is clearly a sum of the first derivatives of $Q_{\mu\nu\lambda\sigma}$ with respect to the components of the k_i , evaluated for certain nonzero values of the k_i ; since p does not enter independently into $Q_{\mu\nu\lambda\sigma}$, but always added to the vectors k_i , we may say that $Q_{\mu\nu\lambda\sigma R}(k_1, k_2, k_3, k_4, p)$ is a sum of the first derivatives of $Q_{\mu\nu\lambda\sigma}$ with respect to the components of p for certain values of p . Therefore, for large $|p|$, the function $Q_{\mu\nu\lambda\sigma R}$ has a factor $|p|^{-1}$ which $Q_{\mu\nu\lambda\sigma}$ does not, and this causes the integral $\int_{-\infty}^{\infty} Q_{\mu\nu\lambda\sigma R} d^4p$ to converge (since $I_{\mu\nu\lambda\sigma}$ diverges logarithmically, $I_{\mu\nu\lambda\sigma R}$ is absolutely convergent).

We shall call $I_{\mu\nu\lambda\sigma R}$ the regularized value of the divergent quantity $I_{\mu\nu\lambda\sigma}$.

We have assumed that parts A, B, C, and D of Fig. 22, are absent, and that the photon lines describe real photons; the above considerations, however, also make it possible to eliminate the divergence due to an irreducible photon-photon scattering part in general, when this part is an internal part of a diagram. To do this we need only subtract from the logarithmically divergent photon-photon scattering part of the matrix element $M(k_1, k_2, k_3, k_4)$, the quantity $M(0, 0, 0, 0)$ which has no physical meaning and should therefore, be discarded. In this way we will obtain the regularized value of $M(k_1, k_2, k_3, k_4)$, given by

$$M_R(k_1, k_2, k_3, k_4) = M(k_1, k_2, k_3, k_4) - M(0, 0, 0, 0), \quad (26.3)$$

which contains no divergence and which should be substituted for $M(k_1, k_2, k_3, k_4)$ in the general expression for the matrix element.

Thus, we see that the requirement of gauge invariance makes it possible to obtain a finite quantity with physical meaning from a divergent matrix element corresponding to an irreducible photon-photon scattering part.

2. Electron and Photon Self-Energies.

Let us now go on to a consideration of the divergences due to electron and photon self-energy parts. We note first that the processes represented by these parts lead to infinite self-energies for the free electron and photon.⁵⁾ In order to show this, let us consider the interaction of the free electron with the zero-point oscillations of the electromagnetic field, that is, with the electromagnetic vacuum state. This process is represented by Fig. 24, diagram 1, which is an electron self-energy part. The difference between this and the general case shown in Fig. 16, diagram 1, is that in this case parts A and B are absent, and the momentum p , since it is the momentum of a real electron, satisfies the relation $p^2 + m^2 = 0$.

From the general rules, we obtain the following expression for the matrix element for this process:

$$S_{p \rightarrow p}^{(2)} = -e^2 \bar{u}(p) \left(\int \gamma_\mu \frac{i \not{p} - m}{p'^2 + m^2} \gamma_\mu \frac{1}{k^2} \delta(p - k - p') \delta(p - k - p') d^4p' d^4k \right) u(p).$$

⁵⁾ This is the reason for calling these "self-energy" parts.

We eliminate one of the δ -functions by integrating over p' , and replace the second one by the four-dimensional integral

$$\delta(q) = \frac{1}{(2\pi)^4} \int e^{iqx} d^4x, \quad q = p - k - p'.$$

Since $q = 0$, the region of integration should be considered bounded. Setting the volume in three-space equal to unity, and the time interval during which the electron interacts with the electromagnetic field equal to Δt , we obtain the following expression for the matrix element $S_{p \rightarrow p'}^{(1)}$:

$$S_{p \rightarrow p'}^{(1)} = \frac{i e^2}{(2\pi)^4} \bar{u}(p) \left(\int \gamma_\mu \frac{i(\hat{p} - \hat{k}) - m}{(p-k)^2 + m^2} \gamma_\mu \frac{d^4k}{k^2} \right) u(p).$$

Now we may consider the interaction of the electron with the zero-point oscillations of the electromagnetic field to lead to an electron energy change. If this change is given by $\Delta \epsilon$, then in a time Δt the electron wave function changes by $(e^{-i\Delta \epsilon \Delta t} - 1) u(p)$. In order to find the matrix element for the transition of the electron from an initial state with energy ϵ before the interaction has taken place to a final state with energy $\epsilon + \Delta \epsilon$, this expression must be multiplied by $\bar{u}^*(p) u(p)$ [we recall that $\bar{u}(p)$ and $u^*(p)$ are related by the normalization condition $\bar{u}^*(p) u(p) = 1$]. Assuming that $\Delta \epsilon \Delta t \ll 1$, we find that the matrix element is given by

$$S_{p \rightarrow p}^{(1)} = -i \Delta \epsilon \Delta t.$$

Comparing this equation with the one derived above, we see that the self-energy of the electron is given by¹⁾

$$\Delta \epsilon = \frac{i e^2}{(2\pi)^4} \bar{u}(p) \left(\int \gamma_\mu \frac{i(\hat{p} - \hat{k}) - m}{(p-k)^2 + m^2} \gamma_\mu \frac{d^4k}{k^2} \right) u(p). \quad (26.4)$$

This expression can be related to an equivalent change of the electron mass Δm :

$$e \Delta \epsilon = m \Delta m.$$

Noting that from the Dirac equation we have

$$\bar{u}(p) u(p) = m u^*(p) u(p) = m,$$

¹⁾ We recall that e is given in Heaviside units.

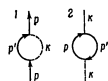


Fig. 24

we obtain the following expression for Δm :

$$\Delta m = \frac{i e^2}{(2\pi)^4} \frac{1}{m} \bar{u} \left(\int \gamma_\mu \frac{i(\hat{p} - \hat{k}) - m}{(p-k)^2 + m^2} \gamma_\mu \frac{d^4k}{k^2} \right) u. \quad (26.4')$$

This mass may be called the electromagnetic mass of the electron. We see that (26.4') diverges linearly for high momenta of the virtual photon, which is in agreement with the previously obtained result on the classification of divergences.²⁾

Let us now consider the interaction between a photon and the electron-positron vacuum state. This process is represented by Fig. 24, diagram 2, that is, by the photon self-energy part. The difference between this diagram and the general one given in Fig. 16, diagram 2 is that in this case $\underline{A}$ and $\underline{B}$ are absent, and since $\underline{k}$ is the momentum of a real photon, it satisfies the relation $k^2 = 0$.

From the general rules for writing down the matrix elements, and by repeating the previous considerations, we obtain the following expression for the photon self-energy $\Delta \epsilon_p$ due to its interaction with the electron-positron vacuum state:

$$\Delta \epsilon_p = i \frac{e^2}{2\pi} \int \text{Spur} \left\{ \gamma_\mu \frac{i\hat{p} - m}{p^2 + m^2} \gamma_\mu \frac{i(\hat{p} - \hat{k}) - m}{(p-k)^2 + m^2} \frac{d^4p}{(2\pi)^4} \right\}. \quad (26.5)$$

This expression is quadratically divergent for large momenta of the virtual electron, which is in agreement with the above results on the classification of divergences.

We thus see that the self-energy parts lead indeed to infinite self-energies of the free electron and photon.

3. Removal of the Divergence due to the Electron Self-Energy.

The occurrence of these infinite energies is a significant defect of the theory. The fact, however, that the self-energy parts of the diagrams are related to infinite self-energies of the free electron and photon makes it possible to develop a general and well-defined method of obtaining finite expressions with physical meaning from divergent matrix elements containing irreducible self-energy parts. This method is based on the fact that the self-energy of the free electron and photon diverge in the existing theory only formally. In other words, when we are dealing with a free electron, its experimentally observable mass m already contains the part due to the electromagnetic mass (if it exists at all). Therefore, we need not take account of the change Δm in the electron mass, and we thus require that this change vanish; that is to say,

$$\Delta \epsilon = 0. \quad (26.6)$$

Similarly, the photon is a particle whose mass and self-energy vanishes. Therefore, we need not take account of the photon self-energy and must set it equal to zero, writing

$$\Delta \epsilon_p = 0. \quad (26.7)$$

²⁾ We note that if the integration in (26.4') is first taken over the angle variables, we obtain an expression which diverges not linearly, but logarithmically (see note on p. 283).

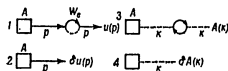


Fig. 25

second-order correction to the scattering process shown in Fig. 25, diagram 2, in which the electron does not interact with the electromagnetic vacuum.

On the basis of the general rules for writing down the matrix elements, we can, according to (25.4), neglect the self-energy part W_e in Fig. 25, diagram 1 and replace the spinor $u(p)$ in diagram 2 by the spinor $\delta u(p)$, according to

$$\delta u(p) = \frac{1}{2} S^p(p) \Sigma(W_e, p) u(p).$$

On the other hand, according to the concepts just described, we should not take account of process 1 at all, since the electromagnetic mass of the electron is already included in the total mass m . Therefore, the spinor $\delta u(p)$ should vanish, and since the operator $\Sigma(W_e, p)$ for the free electron (for which $p^2 = -m^2$) clearly has [see (24.5)] a first-order pole at $\hat{p} = im$, $\Sigma(W_e, p)$ must have a second-order zero at $\hat{p} = im$.

Let us now consider the operator $\Sigma(W_e, p)$ where the momentum p of the virtual electron does not in general satisfy the relation $p^2 + m^2 = 0$. Such an operator, as we know, occurs if an electron self-energy part connects two parts of a diagram (see Fig. 16). We have seen above [see Equation (25.3)] that $\Sigma(W_e, p)$ is an integral of a rational function $R(\hat{t}, \hat{p})$.

$$\Sigma(W_e, p) = \int R(\hat{t}, \hat{p}) dt,$$

where the vector matrix $\hat{p}$ does not enter R independently, but always added to the vector $\hat{t}$ [here t is the momentum of the virtual particle (photon) over which the integration is performed]. This integral is linearly divergent for large $|t|$.

We shall show that $\Sigma(W_e, p)$ can be written in the form

$$\Sigma(W_e, p) = \Sigma_1 + \Sigma_0(\hat{p} - im) + \Sigma_R(W_e, p), \quad (26.8)$$

where Σ_1 and Σ_0 are infinite constants independent of $\hat{p}$ and $\Sigma_R = (\hat{p} - im) \Sigma(\hat{p})$ is a finite operator in which $\Sigma(\hat{p})$ vanishes when $\hat{p} = im$. To do this, let us write $R(\hat{t}, \hat{p})$ in the form

We are then led to an important conclusion with respect to the divergent quantities $\Sigma(W_e, p)$ and $\Pi(W_e, k)$ introduced in (25.2) and (25.6).

Let us first consider the irreducible electron self-energy part represented in Fig. 25, diagram 1. Here A denotes some process giving rise to a free electron with momentum p , but before this electron appears finally as a free particle it interacts with the electromagnetic vacuum state, as indicated by the self-energy part W_e . Clearly Fig. 25, diagram 1, describes a

$$R(\hat{t}, \hat{p}) = R(\hat{t}, 0) + p_\alpha \left(\frac{\partial R}{\partial p_\alpha} \right)_{p=0} + R_\alpha(\hat{t}, \hat{p}).$$

We then obtain the following expression for $\Sigma(W_e, p)$:

$$\Sigma(W_e, p) = A' + B'_\alpha p'_\alpha + \Sigma_0(W_e, p), \quad \Sigma_0(W_e, p) = \int R_0(\hat{t}, p) dt, \quad (26.9)$$

where A' and B'_α are divergent operators independent of p , and $\Sigma_0(W_e, p)$ is an operator containing no divergences. Indeed, R_0 can obviously be written

$$R_0(\hat{t}, \hat{p}) = R(\hat{t}, \hat{p}) - R(\hat{t}, 0) - p_\alpha \left(\frac{\partial R}{\partial p_\alpha} \right)_{p=0} = \frac{1}{2} p_\alpha p_\beta \left(\frac{\partial^2 R}{\partial p_\alpha \partial p_\beta} \right)_{p=0},$$

where the second derivative is evaluated at some point $p = \epsilon \neq 0$. Since $\hat{p}$ enters $R(\hat{t}, \hat{p})$ added to $\hat{t}$, we may state that R_0 is given by the integral of the second derivative of R with respect to $\hat{t}$. It then follows that for large $|t|$ the integrand of $\Sigma_0(W_e, p)$ has a factor of order $|t|^{-3}$ which is not contained in $\Sigma(W_e, p)$, and since $\Sigma(W_e, p)$ diverges linearly, $\Sigma_0(W_e, p)$ converges absolutely.

From invariance considerations it follows that the operator $\Sigma_0(W_e, p)$ should be of the form

$$\Sigma_0(W_e, p) = s_1(W_e, p^2) + s_2(W_e, p^2) \hat{p}, \quad (26.9)$$

where s_1 and s_2 are functions of p^2 . Also on the basis of these considerations, it follows that B'_α is of the form

$$B'_\alpha = \gamma_\alpha B',$$

where B' is an infinite constant.

Equation (26.9*) can be rewritten in the form

This follows from the fact that the polynomial in the denominator of the rational fraction R is of higher degree than that in the numerator.

$$\Sigma_0(W_e, p) = \Sigma_0(W_e, im) + b(p - im) + \Sigma_R(W_e, p),$$

where $\Sigma_R(W_e, p) = (\hat{p} - im) \Sigma(\hat{p})$, and $\Sigma(\hat{p})$ vanishes for $\hat{p} \rightarrow im$. Inserting this expression in (26.9) and introducing the notation

$$A' + \Sigma_0(W_e, im) + im B' = \Sigma_1, \quad B' + b = \Sigma_0,$$

we obtain (26.8), which is a power series expansion of $\Sigma(W_e, p)$ in $\hat{p} - im$.

We have explained above that the divergent operator $\Sigma(W_e, p)$ should have a second-order zero at $\hat{p} = im$. This condition is satisfied by $\Sigma_R(W_e, p)$, but not by the infinite quantities Σ_1 and $\Sigma_0(\hat{p} - im)$ (the first of these operators diverges linearly, and the second logarithmically).

On the basis of these considerations we shall consider $\Sigma_R(W_e, p)$, not $\Sigma(W_e, p)$, to have physical meaning. In other words, we shall drop the infinite terms Σ_1 and $\Sigma_0(\hat{p} - im)$ in (26.8), and shall replace Σ by Σ_R in all matrix elements that contain the linearly divergent operator $\Sigma(W_e, p)$.

We shall call the convergent operator $\Sigma_R(W_e, p)$, which replaces the divergent operator $\Sigma(W_e, p)$ in all matrix elements, the regularized operator Σ .

4. Removal of the Divergence Due to the Photon Self-Energy.

Let us now consider the irreducible photon self-energy part shown in Fig. 25, diagram 3. Here A is some scattering process which gives rise to a photon of momentum k; before this photon becomes a free particle, it interacts with the electron-positron vacuum state, as indicated by the photon self-energy part W_p . It is clear that Fig. 25, diagram 3 gives a second-order correction to the scattering process represented by Fig. 25, diagram 4, in which the emitted photon does not interact with the vacuum. According to (26.8) we may ignore the photon self-energy part W_p , but must replace the operator $\hat{A}(k)$ in the skeleton diagram by

$$\hat{A}(k) = \hat{A}(k) \Pi(W_p, k) \frac{1}{2} D^2(k).$$

On the other hand, on the basis of the previously described considerations, process 3 may be entirely ignored. Therefore, the operator $\hat{A}(k)$ should vanish, and since the function $\Pi(W_p, k)$ has a second-order pole at $k = 0$, the function $\Pi(W_p, k)$ should have a third-order zero at this point.

Let us now consider the function $\Pi(W_p, k)$, where the momentum k of the virtual photon does not in general satisfy the relation $k^2 = 0$. This function, as we know, diverges quadratically and occurs if the photon self-energy part connects two parts A and B of a diagram (see Fig. 16, diagram 2).

According to (25.7) $\Pi(W_p, k)$ can be written as the integral of a certain rational function $Q(\hat{t}, \hat{k})$.

$$\Pi(W_p, k) = \int Q(\hat{t}, \hat{k}) dt,$$

and the vector matrix $\hat{k}$ does not enter Π independently, but added to $\hat{t}$. This integral, as we know, diverges quadratically for large $|\hat{t}|$.

We shall show that $\Pi(W_p, k)$ can be written

$$\Pi(W_p, k) = \Pi_2 + \Pi_0 k^2 + \Pi_R(W_p, k), \quad (26.10)$$

where Π_2 and Π_0 are infinite constants independent of k, and $\Pi_R(W_p, k) = k^2 P(W_p, k)$ is a function which contains no divergences, and in which $P(W_p, k)$ vanishes for $k^2 = 0$.

To do this, let us write $Q(\hat{t}, \hat{k})$ in the form

$$Q(\hat{t}, \hat{k}) = Q(\hat{t}, 0) + k_\alpha \left(\frac{\partial Q}{\partial k_\alpha} \right)_{k=0} + \frac{1}{2} k_\alpha k_\beta \left(\frac{\partial^2 Q}{\partial k_\alpha \partial k_\beta} \right)_{k=0} + Q_0(\hat{t}, \hat{k}),$$

Then for $\Pi(W_p, k)$ we obtain the expression

$$\Pi(W_p, k) = \Pi_2 + k_\alpha C_\alpha + k_\alpha k_\beta \Pi_{\alpha\beta} + \Pi_R(W_p, k), \quad \Pi_R(W_p, k) = \int Q_0(\hat{t}, \hat{k}) dt,$$

where Π_2 and $\Pi_{\alpha\beta}$ are infinite constants independent of k, and Π_R is some finite function of k^2 . Before proving this last equation, we note that $\Pi(W_p, k)$ must be an invariant function of $\hat{k}^2 = k^2$, so that

$$C_\alpha = 0, \quad \Pi_{\alpha\beta} = \Pi_0 \delta_{\alpha\beta}$$

and, therefore,

$$\Pi(W_p, k) = \Pi_2 + k^2 \Pi_0 + \Pi_R(W_p, k).$$

We shall now prove that $\Pi_R(W_P, k)$ is convergent. The function $Q_0(\hat{t}, \hat{k})$ can clearly be written

$$Q_0(\hat{t}, \hat{k}) = Q(\hat{t}, \hat{k}) - Q(\hat{t}, 0) - k_i \left(\frac{\partial Q}{\partial k_i} \right)_{k=0} - \frac{1}{2} k_i k_j \left(\frac{\partial^2 Q}{\partial k_i \partial k_j} \right)_{k=0} = \\ = \frac{1}{6} k_i k_j k_l \left(\frac{\partial^3 Q}{\partial k_i \partial k_j \partial k_l} \right)_{k=0} \neq 0,$$

where the third derivative is evaluated at a point $k = \eta \neq 0$. Since $\hat{k}$ enters $Q(\hat{t}, \hat{k})$ added to $\hat{t}$, we can assert that $\Pi_R(W_P, k)$ is given by the integral of the third derivative of $Q(\hat{t}, \hat{k})$ with respect to $\hat{t}$. It then follows that for large $|\hat{t}|$ the integrand of the expression for $\Pi_R(W_P, k)$ will contain a factor of order $|\hat{t}|^{-1}$ which is not contained in the expression for $\Pi(W_P, k)$; since $\Pi(W_P, k)$ is quadratically divergent, $\Pi_R(W_P, k)$ converges absolutely.

We have explained above that the divergent expression for $\Pi(W_P, k)$ should have a third-order zero at $k = 0$. This condition is satisfied by $\Pi_R(W_P, k)$, but not by the first two terms on the right side of (26.10) (of these, the first diverges quadratically and the second logarithmically).

We shall now assume that the function $\Pi_R(W_P, k)$, not $\Pi(W_P, k)$, has physical meaning. In other words, we shall drop the divergent terms Π_2 and $\Pi_2 k^2$ in expression (26.10), and in all matrix elements containing the quadratically divergent expression $\Pi(W_P, k)$ we shall replace it by the finite function $\Pi_R(W_P, k)$. We shall call $\Pi_R(W_P, k)$ the regularized function $\Pi(W_P, k)$.

5. Removal of the Divergence Due to the Vertex Part.

We have shown how to eliminate divergences from a matrix element containing irreducible self-energy parts.

We shall now show that such removal is possible also for diagrams which contain irreducible vertex parts. If a diagram contains an irreducible vertex part V (see Fig. 20, diagrams 1 and 2), then at vertex 1 of the skeleton diagram (which does not contain this vertex part; see Fig. 21, diagram 3) the operator γ_μ must be replaced by

$$\Delta_\mu = \Delta_\mu(V, p_1, p_2, k) \quad (p_1 - p_2 - k = 0).$$

As we have seen above, this operator is logarithmically divergent for high momenta of the virtual particles. To remove this divergence it is sufficient to note that if $p_1 = p_2 = p_0$, where p_0 is the momentum of the free electron, an operator Δ_μ need not be associated with vertex 1, since in this case the diagram does not represent any real scattering process. In other words, $\Delta_\mu(V, p_0, p_0, 0)$ (if $p_1 = p_2$, then $k = 0$) should be considered zero, and it can, therefore, be subtracted from $\Delta_\mu(V, p_1, p_2, k)$. We shall show that the difference of these operators

$$\Delta_{\mu R}(V, p_1, p_2, k) = \Delta_\mu(V, p_1, p_2, k) - \Delta_\mu(V, p_0, p_0, 0) \quad (26.11)$$

is convergent. To do this, we make use of the integral representation of Δ_μ , namely

$$\Delta_\mu(V, p_1, p_2, k) = \int R(\hat{p}_1, \hat{p}_2, \hat{k}, \hat{t}) dt,$$

where R is some rational function whose integral is logarithmically divergent for large $|\hat{t}|$; here $\hat{t}$ denotes the set of momenta of the virtual particles over which the integration is performed; for a third order irreducible vertex part, R is given by (25.9). The difference $R(\hat{p}_1, \hat{p}_2, \hat{k}, \hat{t}) - R(\hat{p}_0, \hat{p}_0, 0, \hat{t})$ is a sum of products of the components of $\hat{k}$, $\hat{p}_1 - \hat{p}_0$, $\hat{p}_2 - \hat{p}_0$ and the corresponding first derivatives of R evaluated for certain values of these components. Since $\hat{p}_1, \hat{p}_2, \hat{k}$ do not enter R independently, but always added to the vectors $\hat{t}$, it follows that this difference is proportional to the first derivative of R with respect to one of the $\hat{t}$. Therefore, the integrand of the expression for $\Delta_{\mu R}$ for large values of $|\hat{t}|$ will have a factor of order $|\hat{t}|^{-1}$ not contained in the integrand of Δ_μ , and this will guarantee the absolute convergence of $\Delta_{\mu R}$.

We shall now assume that not the operator Δ_μ , but $\Delta_{\mu R}$ has physical meaning; the latter should be substituted for Δ_μ in divergent matrix elements. We shall call $\Delta_{\mu R}$ the regularized operator Δ_μ .

Thus, we see that the four existing types of divergences due to irreducible parts of diagrams can be uniquely eliminated on the basis of simple physical considerations.³⁾

The removal of divergences is accomplished in a well-defined general way, which can be summarized as follows. If $M(p, k)$ is a divergent part of a matrix element due to an irreducible internal part of a diagram (that is, due to self-energy, vertex, and photon-photon scattering parts), and if p and k are the momenta of the electron and photon lines connecting this part to the rest of the diagram, then the divergence in $M(p, k)$ is removed by subtracting from it some of the first terms of its power series expansion in $\hat{p} - \hat{p}_0 = \hat{p} - \hat{p}_0$ or k . The number of terms subtracted should be the least number necessary to guarantee convergence of the remainder $M_R(p, k)$. This remainder is the regularized value of the matrix element $M(p, k)$, and is physically meaningful. As for p and k , these may be either the momenta of internal, or those of external lines of the diagram.

³⁾ When solving concrete problems in practice, the integration in

$$M(p, k) = \int R(\hat{p}, \hat{k}, \hat{t}) dt$$

can first be performed over some finite region. The expression obtained should be regularized by subtracting from it the necessary number of terms of the expansion in powers of $\hat{p} - \hat{p}_0$ and k , and then one may go to the limit of an infinite region of integration.⁴⁾

Since regularization consists of subtracting from a divergent expression several of the first terms of a power series, it follows that the regularization procedure is additive; i.e., if the divergent expression is a sum of several terms, they may be regularized individually. We must bear in mind, however, that if some term is finite, this does not mean that it is identical with the regularized expression.

³⁾ This method of removing divergences is due to Dyson. See the article cited on p. 258.

⁴⁾ In performing the integration over the momenta of the virtual particles one may also introduce various "cut-offs" in the integrand (see Feynman's articles cited on p. 240) which will guarantee the convergence of the integral for large $|\hat{t}|$ (see, in this connection, the regularization method based on auxiliary masses in Section 27).

Regularization can be performed in several stages, dropping those terms at the very start which must vanish in the final result.

6. Removal of Reducible and Overlapping Divergences.

We have considered divergences due to irreducible internal parts of diagrams. The method we have presented, however, is sufficient for removal of divergences due to reducible self-energy, vertex, and photon-photon scattering parts of any complexity.

Let us consider, for instance, a diagram containing the reducible vertex part V_{11} shown in Fig. 26. V_{11} is composed of the irreducible self-energy parts $\bar{W}'_e$, $\bar{W}''_e$, $\bar{W}_e$ and the irreducible vertex part V_3 . It is clear that the part of the matrix element which is associated with V_{11} will diverge not only when integrated over p and k for the diagram as a whole, but also in integrating over p' , p'' , p''' , which refer to each of the component parts $\bar{W}'_e$, $\bar{W}''_e$, $\bar{W}_e$, V_3 , even if the rest of the variables are held fixed.

In this case the divergences may be removed successively according to the above rules, starting with the internal parts $\bar{W}'_e$, $\bar{W}''_e$, $\bar{W}_e$, V_3 , and ending with V_{11} as a whole.

Similarly, divergences due to reducible self-energy parts may be removed by starting with the internal component parts and ending with the self-energy part as a whole. In this case we need only bear in mind that if a vertex part is included at one of the two external vertices of a photon or electron self-energy part, there will simultaneously appear a vertex part at the second external vertex of the self-energy part. Therefore, divergences due to vertex parts in the external vertices of self-energy parts may be called overlapping divergences.

We shall show¹⁾ how to remove overlapping divergences by using as an example the fourth order electron self-energy part W_2 shown in Fig. 18, diagram 5. The operator $\Sigma(W_2, p)$ corresponding to W_2 can be written, according to the general rules for the matrix elements,

$$\Sigma(W_2, p) = \frac{-i}{(2\pi)^4} e^4 \Sigma^{(2)}(W_2, p) = \frac{-i}{(2\pi)^4} e^4 \int d^4t_1 d^4t_2 F_p(p, t_1) G_{\mu\nu}(p, t_1, t_2) H_{\mu\nu}(p, t_2), \quad (26.12)$$

where

$$F_p(p, t_1) = \frac{1}{t_1^2} \gamma_p \frac{t(\hat{p} - \hat{t}_1) - m}{(p - t_1)^2 + m^2}, \quad G_{\mu\nu}(p, t_1, t_2) = \gamma_p \frac{t(\hat{p} - \hat{t}_1 - \hat{t}_2) - m}{(p - t_1 - t_2)^2 + m^2} \gamma_p, \\ H_{\mu\nu}(p, t_2) = \frac{t(\hat{p} - \hat{t}_2) - m}{(p - t_2)^2 + m^2} \gamma_p \frac{1}{t_2^2}$$

¹⁾ A. Salam, Phys. Rev. 82, 217 (1951); 84, 426 (1951).

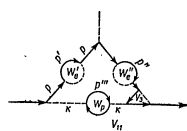


Fig. 26

($\underline{t}_1 = k$, $\underline{t}_2 = k$; see Fig. 18, diagram 5). It is easy to see that $\Sigma(W_2, p)$ is linearly divergent for large $|\underline{t}_1|$, $|\underline{t}_2|$; in integrating over one of the vectors $\underline{t}_1$, $\underline{t}_2$, keeping the other one fixed, we obtain a logarithmically divergent expression, corresponding to the fact that there are vertex parts at $\underline{a}$ and $\underline{b}$. In order to remove the divergence in $\Sigma(W_2, p)$, let us write (26.12) in the form

$$\Sigma^{(2)}(W_2, p) = \int d^4t_1 F_p(p_0, t_1) G_{\mu\nu}(p_0, t_1, 0) \int d^4t_2 H_{\mu\nu}(p, t_2) + \\ + \int d^4t_1 F_p(p, t_1) \int d^4t_2 G_{\mu\nu}(p_0, 0, t_2) H_{\mu\nu}(p_0, t_2) + \int d^4t_1 d^4t_2 R(t_1, t_2, p) d^4t_3, \quad (26.13)$$

where

$$R(t_1, t_2, p) = F_p(p, t_1) G_{\mu\nu}(p, t_1, t_2) H_{\mu\nu}(p, t_2) - \\ - F_p(p_0, t_1) G_{\mu\nu}(p_0, t_1, 0) H_{\mu\nu}(p, t_2) - F_p(p, t_1) G_{\mu\nu}(p_0, 0, t_2) H_{\mu\nu}(p_0, t_2)$$

and p_0 is the momentum of the free electron and thus, satisfies the relation $\underline{p}_0^2 + m^2 = 0$. The first two terms of (26.13) may be ignored, since the integrals

$$\int d^4t_1 F_p(p_0, t_1) G_{\mu\nu}(p_0, t_1, 0) = L_2 \gamma_p, \quad \int d^4t_2 G_{\mu\nu}(p_0, 0, t_2) H_{\mu\nu}(p_0, t_2) = L_2 \gamma_p$$

are divergent expressions contained in the operators $\Lambda_\nu^{(3)}(\underline{p}, \underline{p}_0, \underline{t}_2)$ and $\Lambda_\mu^{(3)}(\underline{p}, \underline{p}_0, \underline{t}_1)$ which characterize the vertex parts. As was explained above, these expressions should be dropped. Therefore, the finite part of $\Sigma^{(2)}(W_2, p)$ is contained in the third term of (26.13). In order to remove the divergence from this term, we shall proceed as with an irreducible electron self-energy part; namely, we shall subtract from R the first two terms of its power series in $\underline{p} - \underline{im}$ or in $\underline{p} - \underline{p}_0$ (where $\underline{p}_0^2 + m^2 = 0$), which is the same thing. As a result we obtain the following expression for the regularized operator $\Sigma^{(4)}(W_2, p)$:

$$\Sigma_R^{(2)}(W_2, p) = \int d^4t_1 d^4t_2 \left\{ R(p, t_1, t_2) - \right. \\ \left. - R(p_0, t_1, t_2) - (p - p_0) \frac{\partial}{\partial p} R(p_0, t_1, t_2) \right\}, \quad (26.14)$$

In order to exhibit the connection between the regularized $\Sigma(W_2, p)$ and the above general method for removing divergences more clearly, let us rewrite $\Sigma_R^{(2)}(W_2, p)$ in the form

$$\Sigma_R(W_2, p) = \Sigma(W_2, p) - e^2 L_0 \int \gamma_1 H_1(p, t_1) dt_1 - e^2 L_0 \int F_2(p, t_1) \gamma_2 dt_1 - D(t_1, t_2), \quad (26.15)$$

where

$$D(t_1, t_2) = e^2 \int [R(p_0, t_1, t_2) - (p - p_0) \frac{\partial}{\partial p} R(p_0, t_1, t_2)] dt_1 dt_2.$$

Here it is clearly seen that the subtraction of the second and third terms removes the divergences due to the vertex parts at a and b , and the subtraction of the fourth term removes the divergence due to W_2 as a whole, from which the internal divergences have already been removed.

The second term in (26.15) is the product of L_0 , which is the divergence due to the vertex part a (that is, the expression which must be subtracted from the matrix element of this vertex part in order to make it finite), and the integral, $\int \gamma_1 H_1(p, t_1) dt_1$, whose integrand contains the function $H_1(p, t_1)$ which is independent of the vertex part at a ; we may say that this integral corresponds to the diagram which is obtained from W_2 when the vertex part at a is removed. Similarly, the third term in (26.15) is the product of L_0 , the divergence due to the vertex part at b , and the integral, $\int F_2(p, t_1) \gamma_2 dt_1$, corresponding to the diagram obtained from W_2 after removal of the vertex part at b . Finally, the third term in (26.15) is the divergence of W_2 as a whole, after the divergences due to the separate internal vertex parts have been removed.

Equation (26.15) can be generalized to the case of arbitrarily complex diagrams, i.e., for diagrams which cannot be broken up into two parts connected by a single line.

In order to make use of the above method of successive removal of the divergences, it is necessary to show also that after regularizing some internal divergence and going on to the integration over the momenta of the virtual particles which are external to this internal part, there arise no new types of divergences which were not listed in Section 26.

For this purpose it is sufficient to show that the modified quantities $\bar{S}^F$, $\bar{D}^F$, Λ_μ have asymptotic behaviors at high momenta which are the same as the quantities S^F , D^F , γ_μ .

Let us first consider $\bar{S}^F(p)$ and $\bar{D}^F(p)$. The differences between these quantities and $S^F(p)$ and $D^F(p)$ can be written in the form of (25.2) and (25.6), where $\Sigma(W_2, p)$ and $\Pi(W_2, p)$ depend on the momentum p and the structure of the self-energy parts W_2 and W_1 . We must find the asymptotic behavior of the regularized Σ and Π for large $|p|$. In doing this we may restrict the considerations to the simplest cases of irreducible self-energy parts W_2 and W_1 , since the divergences are removed, as has been explained above, successively, starting with the internal parts of the diagram and ending with the diagram as a whole.

In the case of irreducible parts W_2 and W_1 , Σ and Π are given by (25.3) and (25.7), which diverge linearly and quadratically, respectively. It will be shown in Section 56 how to calculate this type of integral over some finite and invariant region of the virtual particle momenta.

As can be seen from (56.17) and (56.18), the order of the integral as a function of p (that is, the highest power of p when $|p| \gg m$) is the same as the order of the divergence of the integral, with an accuracy up to a factor $\ln \frac{p^2}{m^2}$. Therefore, up to a factor $\ln \frac{p^2}{m^2}$, the asymptotic behavior of the regularized values of Σ and Π , in the limit when $|p| \gg m$ is similar to $|p|$ and $|p^2|$ respectively. Since $\bar{S}^F$ and $\bar{D}^F$ behave for $|p|$ like $|p|^{-1}$ and $|p^2|^{-1}$, it follows from (25.2) and (25.6) that for $|p| \gg m$ the regularized values $\delta \bar{S}^F$ and $\delta \bar{D}^F$ behave, up to a factor $\ln \frac{p^2}{m^2}$, like $|p|^{-1}$ and $|p^2|^{-1}$, respectively; this behavior is the same as that of $\bar{S}^F$ and $\bar{D}^F$.

Similarly, it is easy to show that the operator Λ_μ given by (25.9), behaves, after regularization, in the limit $|p| \gg m$ like $\ln \frac{p^2}{m^2}$.

In the general case of an arbitrarily complex self-energy or vertex part, the following asymptotic expressions for the regularized values of $\bar{S}^F$, $\bar{D}^F$, Λ_μ are valid:

$$S_R^F(p) \sim S^F(p) f_S\left(\frac{p^2}{m^2}\right), \quad D_R^F(p) \sim D^F(p) f_D\left(\frac{p^2}{m^2}\right), \quad \Lambda_\mu \sim \gamma_\mu f_\mu\left(\frac{p^2}{m^2}\right), \quad |p^2| \gg m^2.$$

Here f_S , f_D , f_μ are dimensionless functions of $\frac{p^2}{m^2}$ which can be written as polynomials in $\ln \frac{p^2}{m^2}$ (in the expression for Λ_μ it is assumed that $|p^2| \sim |p^2| = |p^2| \gg m^2$).

Thus, having removed the divergence in any internal part of a diagram and having gone on to integration over the variables external to it, we arrive at no divergences other than those considered in Section 26.

We note, however, that the above considerations become inapplicable as the order of the diagram approaches infinity, since in this case the polynomials in $\ln \frac{p^2}{m^2}$ can become infinite series, which can change the asymptotic behavior of the functions f_S , f_D , f_μ .

$$\text{If, for instance, it is found that these functions are given by the series } f = \sum_{n=0}^{\infty} \frac{e^{2n} (\ln \frac{p^2}{m^2})^{2n}}{n!} = \left(\frac{p^2}{m^2}\right)^e,$$

the character of the divergence changes greatly on going from the internal parts of the diagram to the diagram as a whole.

§ 27. Mass and Charge Renormalization.

1. The Renormalization Concept.

We shall now go on to a discussion of the physical concepts at the basis of the above regularization method.

In studying the divergence due to the photon self-energy part, we have shown that it leads to infinite photon self-energy, and on this basis we have replaced the divergent operator $\delta \hat{A}(k)$ corresponding to the dotted line of Fig. 25, diagram 4 by zero. If, however, the dotted line of Fig. 25, diagram 4 does not represent a photon, but some given external electromagnetic field, then the divergent operator $\delta \hat{A}(k)$ cannot be replaced by zero.

In this case we can obtain a finite expression having physical meaning from $\delta\hat{A}(k)$ according to the above rules; this will then describe vacuum polarization.

It is known that the D'Alembertian, when applied to a potential $\frac{1}{\mu}(x)$ describing a given external field, leads to the current $\frac{1}{\mu}(x)$ which gives rise to the field. Thus, $-\square\delta A_{\mu}(x)$ is a correction to the external current $\frac{1}{\mu}(x)$, which is due to the interaction of this current with the zero-point oscillations of the electron-positron field. In Section 43 it will be shown that

$$-\square\delta A_{\mu}(x) = a_1 I_{\mu}^{(0)}(x) + I_{\mu}^{(1)}(x),$$

where a_1 is a quadratically infinite constant, and $I_{\mu}^{(1)}(x)$ is finite. If we remove the divergence from δA_{μ} , then the regularized operator $\delta A_{\mu R}$ satisfies the relation

$$-\square\delta A_{\mu R}(x) = I_{\mu}^{(1)}(x). \quad (27.1)$$

We may, therefore, say that the regularization procedure removes an infinite term $a_1 I_{\mu}^{(0)}(x)$ from this correction to the external current; this infinite term is proportional to the original external current $\frac{1}{\mu}(x)$.

In particular, if we have a charge e (for instance an electron), then its interaction with the zero-point oscillations of the electron-positron field causes this charge to change by an infinite amount proportional to e . The regularization procedure consists of not taking account of this addition, assuming that it has no physical meaning. It may be said that the addition to the charge of the electron cannot be separated from the charge itself, and that the regularization process reduces essentially to renormalization of the charge; the sum of the hypothetical electron charge which does not interact with the zero-point oscillations of the electron-positron field, and the infinite correction, which modern theory gives to this charge in view of this interaction, is actually finite and is the total experimentally observed electron charge.

A similar situation arises in removing divergences due to the electron self-energy part. We have seen above that these divergences lead to an infinite electromagnetic mass of the electron, that is, to an infinite correction to the mass of the electron due to its interaction with the zero-point oscillations of the electromagnetic field. The regularization method consists of ignoring this correction, assuming that it cannot be separated from the total electron mass. Thus, we may say that the regularization process reduces to a renormalization of the electron mass: the sum of the mass of the "bare" hypothetical electron, which does not interact with the zero-point oscillations of the electromagnetic field, and its infinite electromagnetic mass, which the theory predicts as a result of this interaction, is actually finite and is the total experimentally observable electron mass.

Thus, the physical concepts underlying the above regularization method are essentially contained in the renormalization of the constants m and e .

We shall now attempt a more rigorous formulation of the renormalization procedures for the mass and charge of the electron.¹⁾ Let us first recall that in quantum electrodynamics we have to deal with six types of

infinite constants related to the irreducible diagrams for the electron and photon self-energy, the vertex parts, and the photon-photon scattering parts. These constants Σ_1 , Σ_2 , Π_0 , Π_1 , Π_2 , M_0 enter the operators $\Sigma(W_e, p)$, $\Pi(W_p, k)$, Λ_{μ} and the photon-photon scattering matrix element in the following way:

$$\left. \begin{aligned} \Sigma(W_e, p) &= \Sigma_1 + \Sigma_2(\hat{p} - \hat{p}_0) + \Sigma_R(W_e, p), \\ \Pi(W_p, k) &= \Pi_0 + \Pi_1 k^2 + \Pi_R(W_p, k), \\ \Lambda_{\mu}(V, p_1, p_2, k) &= L_0 \gamma_{\mu} + \Lambda_{\mu R}(V, p_1, p_2, k), \\ M(k_1, k_2, k_3, k_4) &= M_0 + M_R(k_1, k_2, k_3, k_4). \end{aligned} \right\} \quad (27.2)$$

We have seen above that Σ_2 , Π_0 , L_0 , M_0 are logarithmically divergent, Σ_1 is linearly divergent, and Π_1 is quadratically divergent. As for the infinite constants M_0 and Π_0 , they are independent of the photon momenta and can thus be simply dropped from considerations of gauge invariance. We may thus say that there are only four types of infinite constants, namely Σ_1 , Σ_2 , Π_1 , L_0 .

We shall now show that two of these constants, namely the linearly divergent one Σ_1 and the logarithmically divergent one Π_1 , can be eliminated if the form of the interaction between the electron-positron and the electromagnetic fields is altered slightly.

Thus far, we have used the following expression for the interaction energy density:

$$V(x) = -j_{\mu}(x) A_{\mu}(x). \quad (27.3)$$

Let us now add to this expression the two terms $-\delta\bar{m}\bar{\psi}\psi$ and $-\frac{1}{4}\delta f F_{\mu\nu}^2$, where $F_{\mu\nu}$ is the electromagnetic field tensor, and $\delta\bar{m}$ and δf are certain constants (infinite, but independent of ψ and $F_{\mu\nu}$); this means that we shall consider the interaction energy density given by

$$V^*(x) = -j_{\mu}(x) A_{\mu}(x) - \delta\bar{m}\bar{\psi}(x)\psi(x) - \frac{1}{4}\delta f F_{\mu\nu}^2(x). \quad (27.4)$$

If this expression is inserted for $V(x)$ into the general formula (22.2) for the S matrix, the form of the S matrix is changed, as is that of the operators $\Sigma(W_e, p)$ and $\Pi(W_p, k)$. It is clear that the second term in (27.4) will affect electron transitions, and since it does not contain the electron momentum, appropriate choice of $\delta\bar{m}$ will make it possible to remove the divergent term Σ_1 in the expression for $\Sigma(W_e, p)$. Similarly, the third term in (27.4) will affect photon transitions, and since it is proportional to the square of the field tensor, or the square of the photon momentum, appropriate choice of δf will make it possible to remove the divergent term $\Pi_1 k^2$ in the expression for $\Pi(W_p, k)$.

Thus, the addition of the two terms $-\delta\bar{m}\bar{\psi}\psi$ and $-\frac{1}{4}\delta f F_{\mu\nu}^2$ to the expression for the interaction energy density, makes it possible to eliminate the two divergent constants Σ_1 and Π_1 from the theory. But the physical meaning of such an addition to the interaction term is that the electron mass and the electromagnetic field are being renormalized. Indeed, the Lagrangian of the free electron-positron field contains the term

¹⁾ S. Gupta, Proc. Phys. Soc. (London) A 64, 426 (1951); F. Dyson, Phys. Rev. 83, 608 (1951).

$-\frac{m}{2} \bar{\psi} \psi$, and, therefore, the additional term $-\delta m \bar{\psi} \psi$ can be thought of as the renormalization of the electron mass. The Lagrangian of the free electromagnetic field is equal to $-\frac{1}{4} F_{\mu\nu} F_{\mu\nu}$, and therefore, the correction $-\frac{1}{4} \delta F_{\mu\nu} F_{\mu\nu}$ causes the renormalization of the electromagnetic field, that is, replacement of $F_{\mu\nu}$ by the expression

$$F_{\mu\nu} = F_{\mu\nu} (1 + \delta f)^{1/2}.$$

This field renormalization can also be thought of as renormalizing the electron charge, that is to say, replacing the charge e by the expression

$$e^* = e (1 + \delta f)^{-1/2}.$$

We thus see that the divergent terms Σ_1 and Π_0 are effectively removed by renormalizing the electron's mass and charge. (We note that the removal of Π_0 and M_0 is equivalent to renormalizing the photon mass, since gauge invariance follows from the fact that the photon mass vanishes).

2. The Relation Between the Divergences Due to the Vertex Part and the Electron Self-Energy Part.

After renormalizing the charge and mass of the electron, we are left with the two infinite constants Σ_0 and Π_0 . We shall now show that in any given matrix element they lead to expressions which cancel each other. We first make note of the identity

$$\frac{-1}{(2\pi)^4} \frac{\partial S^F(p)}{\partial p_\mu} = \frac{1}{2} S^F(p) \gamma_\mu \frac{1}{2} S^F(p). \quad (27.5)$$

On the basis of this identity, it can be shown that¹⁾

$$\frac{\partial \Sigma^{(2)}(W, p)}{\partial p_\mu} = -(2\pi)^4 \Delta_\mu^{(0)}(V, p, p), \quad (27.6)$$

where $\Sigma^{(2)}(W, p)$ and $\Delta_\mu^{(0)}(V, p, p)$ are operators referring to the electron self-energy part W and the vertex part V shown in Fig. 27 (see Equations (26.3) and (26.9)).

Equation (27.6) has a simple graphical interpretation. Since $\frac{\partial S^F(p)}{\partial p_\mu}$ is represented by an electron line, we

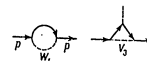


Fig. 27

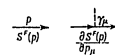


Fig. 28

may say that $\frac{\partial S^F(p)}{\partial p_\mu}$ should be represented by a vertex with one photon and two electron lines (the electron momenta are equal; see Fig. 28). It is then simple to establish the following general rule for the graphic representation of the derivative of a matrix element with respect to an electron momentum. In differentiating any matrix element by an electron momentum, we obtain a series of expressions which can be represented by the set of diagrams in each of which an electron line of the original diagram (corresponding to the original matrix element) is replaced by a vertex with one photon and two electron lines. For instance, the matrix element $\Sigma(W, p)$ represented by the diagram W of Fig. 29a has a derivative $\frac{\partial \Sigma(W, p)}{\partial p_\mu}$ consisting of the three terms whose diagrams are shown in Fig. 29b.

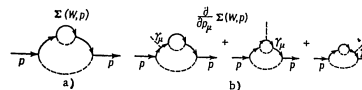


Fig. 29

If we use these diagrams to represent the derivatives of matrix elements, it is easy to show that

$$-\frac{e}{(2\pi)^4} \frac{\partial \Sigma^*}{\partial p_\mu} = \Delta_\mu^*, \quad (27.7)$$

where Σ^* is the sum of the expressions $\Sigma(W, p)$ for all the electron self-energy proper parts W and Δ_μ^* is the sum of the expressions $\Delta_\mu(V, p, p)$ for all the proper vertex parts V .

We note that the operator Σ^* is related to the modified function $S^{F*} = S^F + \delta S^F$ which accounts for all the radiative corrections by the expression²⁾

¹⁾ A part of a diagram is called proper if it cannot be separated into two disconnected parts by the omission of a single line.

²⁾ See Section 45, Equation (45.20), and Dyson's article cited on p. 258.

¹⁾ J. Ward, Phys. Rev. 73, 182 (1950).

$$\frac{1}{2} S^P = \frac{1}{2} S^F + \frac{1}{2} S^F \Sigma^* \frac{1}{2} S^P. \quad (27.8)$$

Let us now illustrate the cancellation of divergences due to vertex parts and electron self-energy parts by two examples (we shall assume that the charge and mass renormalization have been previously performed).

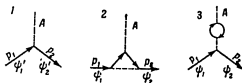


Fig. 30

Let us first consider the radiative corrections to electron scattering in an external field up to third-order terms. The diagrams representing these processes are shown in Fig. 30. Here diagrams 2 and 3 show third-order effects, and diagram 1 illustrates the nonzero matrix element in the first approximation. We should, however, bear in mind that this diagram also describes a third-order effect, since according to (25.10), accounting for the radiative corrections reduces not only

to replacing S^E and D^E by the modified functions $S^E = S^E + \delta S^E$ and $D^E = D^E + \delta D^E$, but also assumes

that the electron wave function operator ψ is replaced by some modified operator $\psi' = \psi + \delta\psi$. (This replacement is equivalent to considering, in addition to diagram 1 with the unmodified ψ , also diagrams in which the electron lines p_1 and p_2 have self-energy parts Σ . See Fig. 27).

The replacement of ψ by ψ' can be performed in the following way. Since $\frac{1}{2} \epsilon S^E$ is the vacuum expectation value of $\bar{\psi}(\psi\bar{\psi})$,

$$\langle P(\bar{\psi}_e(x) \bar{\psi}_p(x')) \rangle_0 = \frac{1}{2} \epsilon (x', x) S^E(x - x'), \quad (27.9)$$

we may say that $\frac{1}{2} \epsilon S^E$ is the vacuum expectation value of $\bar{\psi}(\psi\bar{\psi})$, given by

$$\langle P(\bar{\psi}_e(x) \bar{\psi}_p(x')) \rangle_0 = \frac{1}{2} \epsilon (x', x) S^E(x - x'). \quad (27.9')$$

On the other hand, in the approximation we are now considering [see (25.2)],

$$\frac{1}{2} S^P(p) = \frac{1}{2} S^F(p) + \frac{1}{2} \delta S^P(p) = \frac{1}{2} S^F(p) + \frac{1}{2} S^F(p) \Sigma(W_1, p) \frac{1}{2} S^F(p), \quad (27.10)$$

where the electron self-energy part W_1 is that indicated in Fig. 27. Since

$$\Sigma(W_1, p) = \Sigma_1 + (p - p_0) \left(\frac{\partial \Sigma}{\partial p} \right)_{p=p_0} + \Sigma_R(W_1, p)$$

and

$$\left(\frac{\partial \Sigma}{\partial p} \right)_{p=p_0} = -\frac{(2\pi)^4}{e} \Delta_F(V, p_0, p_0) = -(2\pi)^4 e^2 L_0^0 \gamma_p,$$

we have

$$\Sigma(W_1, p) = \Sigma_1 - (2\pi)^4 e^2 L_0^0 (p - p_0) \gamma_p + \Sigma_R(W_1, p) = \Sigma_1 - (2\pi)^4 e^2 L_0^0 (\hat{p} - im) + \Sigma_R(W_1, p). \quad (27.11)$$

Renormalizing the electron mass we can drop the linearly divergent constant Σ_1 from (27.11), and use the following expression for $\Sigma(W_1, p)$:

$$\Sigma(W_1, p) = -(2\pi)^4 e^2 L_0^0 (\hat{p} - im) + \Sigma_R(W_1, p). \quad (27.11')$$

Noting that $\frac{1}{2} S^F(p) = \frac{1}{(2\pi)^4} (\hat{p} - im)^{-1}$, and inserting (27.11) into (27.10), we obtain

$$\frac{1}{2} S^P(p) = \frac{1}{2} S^F(p) - e^2 L_0^0 \frac{1}{2} S^F(p) + \frac{1}{2} S^F(p) \Sigma_R(W_1, p) \frac{1}{2} S^F(p). \quad (27.12)$$

For the free electron the last term in (27.12), as we know, vanishes so that

$$S^P(p) = (1 - e^2 L_0^0) S^F(p), \quad p = p_0. \quad (27.13)$$

From this and from (27.9') it follows that

$$\psi' = (1 - e^2 L_0^0)^{1/2} \psi \approx \left(1 - \frac{1}{2} e^2 L_0^0 \right) \psi. \quad (27.14)$$

(The last equation is valid if $L_0^{(s)}$ is only formally infinite; actually it gives a small correction proportional to e^2).

Let us now determine the sum of the matrix elements corresponding to diagrams 1, 2, and 3 of Fig. 30. This sum can be written in the form

$$S_{f \rightarrow f}^{(0)} = e \bar{u}_2 \hat{a}(q) u_1 + e^2 \bar{u}_2 \hat{a}(q) u_1 + e^2 \bar{u}_2 \hat{a}(q) u_1 + \mathcal{M}_0, \quad q = p_2 - p_1, \quad (27.15)$$

where $\mathcal{M}_0$ is the matrix element corresponding to diagram 3. In renormalizing the electron charge we will remove the divergence in $\mathcal{M}_0$. We shall therefore consider $\mathcal{M}_0$ already finite.

We must now show that the divergences in the first two terms of (27.15) cancel each other. This, however, follows directly from (27.14) and the form of $\Delta_\mu^{(s)}$, which can be written

$$\Delta_\mu^{(0)}(V, p_1, p_2) = L_0^{(0)} \gamma_\mu + \Delta_{\mu R}^{(0)}(V, p_1, p_2). \quad (27.15')$$

Indeed, by inserting this expression into (27.14) and (27.15) we obtain a finite expression, namely

$$S_{f \rightarrow f}^{(0)} = e \bar{u}_2 \hat{a}(q) u_1 (1 - e^2 L_0^{(0)}) + e^2 \bar{u}_2 \hat{a}(q) u_1 + e^2 \bar{u}_2 \hat{a}(q) u_1 + \mathcal{M}_{2R} = e \bar{u}_2 \hat{a}(q) u_1 + e^2 \bar{u}_2 \hat{a}(q) u_1 + \mathcal{M}_{2R}. \quad (27.16)$$

Thus, we see that the divergence in $S_{f \rightarrow f}^{(s)}$ due to the vertex part cancels that which remains in the electron self-energy part after mass renormalization (this latter divergence is contained in the modified free-electron operator). This can also be put differently; the divergence due to the vertex part is removed after the renormalization of the free-electron wave function. As a second example of the way the vertex part divergence cancels with that remaining in the electron self-energy part after mass renormalization, let us consider the polarization of the vacuum up to terms in e^4 . The diagrams describing this process are shown in Fig. 31.

The functions $\Pi(W, k)$ corresponding to these diagrams (see Equation (25.6)) shall be denoted by $\Pi_a(k)$, $\Pi_b(k)$, and $\Pi_c(k)$. The total function $\Pi(k)$ corresponding to vacuum polarization is clearly given by

$$\Pi(k) = \Pi_a(k) + 2\Pi_b(k) + \Pi_c(k). \quad (27.17)$$

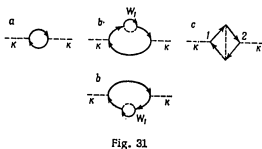


Fig. 31

$\bar{u}$ is the spinor amplitude of ψ .

In order to remove the divergences in $\Pi(k)$, let us first renormalize the electron charge. Then $\Pi_a(k)$ becomes finite. As for Π_b and Π_c , they contain no principal divergences (that is, divergences related to the diagram as a whole), although they do contain divergences related to the internal parts of the diagram ($\Pi_b(k)$ contains the divergence related to the internal electron self-energy part $\mathcal{W}_1$ and $\Pi_c(k)$ contains a divergence related to the two vertex parts at 1 and 2). We shall now show that these divergences due to the vertex and electron self-energy parts cancel in the function $\Pi(k)$ after mass renormalization. We shall do this by using Equation (25.2) for $\frac{1}{2} \delta S^F(p)$. According to (25.10) and (25.2), $\Pi_b(k)$ can be written

$$\Pi_b(k) = -e^2 \int \gamma_\mu \frac{1}{2} S^F(p) \Sigma(W_1, p) \frac{1}{2} S^F(p) \gamma_\mu \frac{1}{2} S^F(p+k) d^4p. \quad (27.18)$$

Let us now perform the mass renormalization. Then $\Sigma(W, p)$ is written in the form given by (27.11'). Inserting (27.11') into (27.18) and bearing in mind the definition of $\Pi_a(k)$ [see (25.7)], we obtain

$$\Pi_b(k) = -e^2 L_0^{(0)} \Pi_a(k) + \Pi_{bR}(k),$$

where $\Pi_{bR}(k)$ is finite (we have made use of the fact that $\frac{1}{2} S^F(p) = \frac{1}{(2\pi)^4} (p - im)^{-1}$). Similarly, using Equations (25.10) and (27.15'), $\Pi_c(k)$ can be given by

$$\Pi_c(k) = 2e^2 L_0^{(0)} \Pi_a(k) + \Pi_{cR}(k),$$

where $\Pi_{cR}(k)$ remains finite. Inserting these expressions for $\Pi_b(k)$ and $\Pi_c(k)$ into (27.17), $\Pi(k)$ is given by

$$\Pi(k) = \Pi_a(k) - 2e^2 L_0^{(0)} \Pi_a(k) + 2\Pi_{bR}(k) + 2e^2 L_0^{(0)} \Pi_a(k) + \Pi_{cR}(k) = \Pi_{bR}(k) + 2\Pi_{bR}(k) + \Pi_{cR}(k), \quad (27.19)$$

where $\Pi_{aR} = \Pi_a$ (because of the charge renormalization this function does not contain any divergences).

Thus, as was asserted above, after mass and charge renormalization we obtain cancellation of the divergences due to the vertex and electron self-energy parts. We emphasize again that if a diagram contains external electron lines, such cancellation of divergences is obtained only after the free-electron wave functions for these external lines have been renormalized.

3. Removal of Divergences by Means of Auxiliary Masses.

In addition to the above way of removing the divergences from the S matrix, there exists another method

of regularizing the divergent matrix elements, which is based on a purely formal introduction of auxiliary masses.³⁾

This method consists of the following. The commutators and anticommutators of the electromagnetic and electron-positron fields, as well as the vacuum expectation values of products of their components can, as we know, be written in terms of the two functions $\Delta(x)$ and $\Delta^{(1)}(x)$ which have singularities on the light cone $x^2 = x^2 - t^2 = 0$. One could expect that the divergences in modern theory are due essentially to the singularities in $\Delta(x)$ and $\Delta^{(1)}(x)$.

The sharpest of these singularities can be removed by replacing $\Delta(x)$ and $\Delta^{(1)}(x)$ by the regularized invariant functions $\Delta_R(x)$ and $\Delta_R^{(1)}(x)$, defined by

$$\Delta_R(x) = \sum_i c_i \Delta(x, M_i), \quad \Delta_R^{(1)}(x) = \sum_i c_i \Delta^{(1)}(x, M_i), \quad (27.20)$$

where $\Delta(x, M_i)$ and $\Delta^{(1)}(x, M_i)$ are functions similar to $\Delta(x)$ and $\Delta^{(1)}(x)$ except that the electron mass m is replaced by some arbitrary mass M_i , and the c_i are constants satisfying

$$\sum_i c_i = 0, \quad \sum_i c_i M_i^2 = 0, \quad c_0 = 1, \quad M_0 = m. \quad (27.21)$$

Since for small x^2 [see (18.50)]

$$\Delta(x) = \frac{i}{4\pi} \left\{ \delta(x^2) + \frac{m^2}{4} \theta(x^2) \right\},$$

$$\Delta^{(1)}(x) = \frac{1}{4\pi} \left\{ \frac{2}{x^2} + m^2 \ln \frac{1}{2} (m^2 |x^2|)^{1/2} - \frac{m^2}{2} \right\},$$

It can be shown rather simply that conditions (27.21) actually remove the $\frac{1}{x^2}$ and $\delta(x^2)$ types of singularities in $\Delta(x)$ and $\Delta^{(1)}(x)$. The regularized function $\Delta_R(x)$ vanishes for $x^2 = 0$, whereas, also on the light cone, $\Delta_R^{(1)}(x)$ is given by

$$\Delta_R^{(1)}(0) = \frac{1}{4\pi} \sum_i c_i M_i^2 \ln M_i. \quad (27.21')$$

³⁾ Pauli and Villars, *Rev. Mod. Phys.* 21, 434 (1949) [a Russian translation of this article can be found in the book, *Atomic Electron Level Shifts* (Foreign Lit. Press, 1951) p. 139].

This method for constructing the regularized functions $\Delta_R(x)$ and $\Delta_R^{(1)}(x)$ can be used to remove the sharpest divergences in the matrix elements by forming from them combinations similar to (27.20).

If we are given some divergent matrix element

$$S(\Delta(x, m), \Delta^{(1)}(x, m); \Delta(x, 0), \Delta^{(1)}(x, 0))$$

which is a function of $\Delta(x, m)$, $\Delta^{(1)}(x, m)$, $\Delta(x, 0)$ and $\Delta^{(1)}(x, 0)$, it should be replaced by the regularized expression

$$S_R = \lim_{M_1, M_2, \dots \rightarrow \infty} \sum_i c_i S(\Delta(x, \sqrt{m^2 + M_i^2}), \Delta^{(1)}(x, \sqrt{m^2 + M_i^2}); \Delta(x, M_i), \Delta^{(1)}(x, M_i)), \quad (27.22)$$

where $\Delta(x, \sqrt{m^2 + M_i^2})$ and $\Delta^{(1)}(x, \sqrt{m^2 + M_i^2})$ are substituted into $S(\Delta(x, m), \Delta^{(1)}(x, m); \Delta(x, 0), \Delta^{(1)}(x, 0))$ in place of the electron functions $\Delta(x, m)$ and $\Delta^{(1)}(x, m)$, and $\Delta(x, M_i)$ and $\Delta^{(1)}(x, M_i)$ are substituted for the photon functions $\Delta(x, 0)$ and $\Delta^{(1)}(x, 0)$. [Note that $\Delta_R \neq S(\Delta_R, \Delta_R^{(1)}, \dots)$.]

Regularization performed in this way, however, is not entirely unique, since for $x^2 = 0$ the regularized function $\Delta_R^{(1)}(x)$ has an undefined value as given by (27.21'). In particular, the electromagnetic mass of the electron Δm and the correction to the electron charge Δe which is due to vacuum polarization are not well defined. Depending on the choice of the M_i and the c_i , the quantities Δm and Δe can take on many different values, including zero.⁴⁾ Therefore, final regularization requires additional considerations similar to those used in obtaining Δ_R , $\Delta_R^{(1)}$ and $\Delta_R^{(2)}$.

These two methods of regularization lead to the same results in many concrete problems, although there does not exist a general proof of their equivalence. Since the auxiliary mass method does not lead to unique removal of all divergences, we shall in the future eliminate divergences directly from the examination of the diagrams, as has been explained previously.

⁴⁾ If we use just one auxiliary mass M_1 (with $c_0 = 1$, $c_1 = -1$), we obtain the following expressions for Δm and Δe :

$$\Delta m = \frac{3\alpha m}{2\pi} \left(\ln \frac{M_1}{m} + \frac{1}{4} \right), \quad \Delta e = -\frac{\alpha e}{3\pi} \ln \frac{M_1}{m},$$

where α is the fine structure constant.

4. On the Divergence of the Renormalized Series for the S Matrix.

The renormalization methods we have described do not remove the divergences in the S matrix as a whole, but only in its separate matrix elements, or in other words in the individual terms of the expansion of the S matrix as a power series in the electron charge. Therefore, the question arises whether or not this power series converges after its individual terms have been renormalized. If it were to converge, one could consider the sum to be the actual S matrix. There are, however, reasons for supposing the series in powers of e to diverge even after renormalization.¹⁾ These reasons are contained in the following remarks.

As will be shown in Section 35, the interaction of two electrons without the emission and absorption of photons can be described by the interaction function $e^2 D(x_1 - x_2)$, where x_1 and x_2 are the position four-vectors of the electrons. Let us use this function to calculate some physical quantity $F(e^2)$ which can be written in the form of a power series in e^2 :

$$F(e^2) = a_0 + e^2 a_2 + e^4 a_4 + \dots \quad (27.23)$$

Let us assume that this series, whose individual terms have been renormalized by the previously described rules, converges for some value of e^2 . This means that $F(e^2)$ is an analytic function of e^2 at $e^2 = 0$. In that case, for a sufficiently small value of the charge, $F(-e^2)$ will also be an analytic function which can be written as a convergent power series. $F(-e^2)$ has a simple physical interpretation. This function would be the quantity F under consideration if the interaction of two charges were described by $-e^2 D(x)$, rather than $+e^2 D(x)$, or in other words, if like charges attracted instead of repelled.

We note that if this were the case, the usual definition of the vacuum would not correspond to the lowest energy. Indeed, let us suppose that N electron-positron pairs are created, and that all the electrons are located in one region of space and all the positrons in another. If these regions are sufficiently small and sufficiently far apart, then for large N the negative Coulomb energy which pulls together the like particles can be greater than the rest energy and the kinetic energy of the particles (this can be accomplished with distances between the particles large enough for simple classical energy calculations to be valid). We shall call such a state briefly "pathological".²⁾

Let us now assume that the interaction between charges is given by the function $-e^2 D(x_1 - x_2)$, and let us consider an ordinary state involving several particles. This state then is separated by a potential barrier from the "pathological" state with the same energy, the height of this barrier being the energy necessary for the creation of N pairs, or the rest energy of $2N$ particles. Nevertheless, because of the quantum mechanical tunnel effect, there is a finite probability that the system will spontaneously pass through the barrier, going over from the ordinary to the "pathological" state. This means that with this interaction function every physical state is unstable with respect to the spontaneous creation of large numbers of particles. Further, the "pathological" state into which the system goes over, is not stationary, since more and more particles will be created; it may be said that there takes place a disintegration of the vacuum. In view of these effects, one cannot assume that quantum electrodynamics with an interaction function of the form $-e^2 D(x_1 - x_2)$ should lead to well-defined analytic functions. In other words, $F(-e^2)$ cannot be analytic and, therefore, (27.23) may be divergent for $e^2 > 0$.

It would seem that the divergence of this series is associated with processes in which large numbers of particles are involved. Therefore, the divergence of (27.23) becomes noticeable if we consider terms of very

high order. The terms of this series first decrease, and then start to increase without limit, the increase beginning when the number of terms is about 137. The first few terms of (27.23), which is what is ordinarily used in calculating concrete physical processes, give an asymptotic representation of the desired function $F(e^2)$, and as is shown by comparison with experiment, this is a good approximation to the function. Since, however, $F(e^2)$ is not an analytic function of e^2 , this asymptotic series cannot serve for a complete and unique definition of this function. Since in its present state of development quantum electrodynamics makes it possible to calculate only the coefficients $a_0, a_2, \dots$, and since in view of the nonanalytic character of $F(e^2)$ these coefficients are insufficient for its complete determination, quantum electrodynamics cannot be considered a complete theory.

It should also be pointed out that the interaction of the electromagnetic field with meson fields gives corrections even to the first terms of (27.23). Indeed, the electromagnetic field can create and then annihilate a virtual meson pair. The probability for this process, relative to the creation and annihilation of a virtual electron-positron pair, is $(\frac{m}{\mu})^2$ (where μ is the meson mass, and m is the electron mass). Since $\frac{m}{\mu} \sim \frac{1}{200}$, these effects need be accounted for only when the approximation is two or three orders of magnitude higher than in calculating similar effects involving the electron-positron field. Therefore, it would seem that all the difficulties of modern quantum electrodynamics could be eliminated in a more complete theory which would take into account the diversity of physical interactions, including those with meson fields, instead of merely that set of phenomena which involves only the electromagnetic and electron-positron fields.

§ 28. The Probability for Various Processes.

1. General Probability Formula.

Having used the rules of Section 24 to determine the matrix element for some transition $i \rightarrow f$, the probability W for this process is given by

$$W = |S_{i \rightarrow f}^{(n)}|^2. \quad (28.1)$$

If only free electrons and photons participate in the process, then clearly the matrix element can be written

$$S_{i \rightarrow f}^{(n)} = M(p_i, p_f) \delta(\Sigma p_i - \Sigma p_f), \quad (28.1')$$

where $\delta(2p)$ is the four-dimensional δ -function of the difference between the initial (p_i) and final (p_f) four-momenta of the particles. Therefore,

$$W = |M(p_i, p_f)|^2 \delta^4(\Sigma p_i - \Sigma p_f).$$

We shall replace one of the δ -functions in this expression by the integral

$$\delta(q) = \frac{1}{(2\pi)^4} \int e^{iqx} d^4x, \quad q = \Sigma p_i - \Sigma p_f.$$

¹⁾ F. Dyson, Phys. Rev. 85, 631 (1952). There is no rigorous proof of the divergence for the case of electrodynamics. It has, however, been shown for a simpler type of interaction. See Thirring, Helv. Phys. Acta 26, 33 (1953).

²⁾ The possibility of "pathological" states is related to the fact that the interaction function $-e^2 D(x_1 - x_2)$ is essentially equivalent to using a non-Hermitian Hamiltonian with an imaginary electric charge.

and since $q = 0$, the region of integration should be considered bounded. Assuming that the volume in three-space in which the particle wave functions are normalized is unity, and denoting the time interval by t , we obtain the following expression for W :

$$W = \frac{1}{(2\pi)^4} |M(p_i, p_f)|^2 \delta(\Sigma p_i - \Sigma p_f) t.$$

Since the probability for the transition $i \rightarrow f$ is proportional to the time, we may speak of the transition probability per unit time, namely

$$w_i = \frac{1}{(2\pi)^4} |M(p_i, p_f)|^2 \delta(\Sigma p_i - \Sigma p_f). \quad (28.2)$$

In the processes we are considering, the initial and final states belong to the continuous spectrum. It is, therefore, of interest to find the probability that the three-dimensional momenta $\underline{p}_f$ in the final state are in the interval d^3p_f . To find this probability, which we shall denote $d\omega$, the probability per unit time w_i must be multiplied by $\Pi(2\pi)^{-3} d^3p_f$, where $(2\pi)^{-3} d^3p_f$, as is well known, is the number of states with a given orientation of spin (in the case of electrons) or a given polarization (in the case of photons) for which $\underline{p}_f$ is in the interval d^3p_f (the normalizing volume is assumed unity). Thus,

$$d\omega = \frac{1}{(2\pi)^4} |M(p_i, p_f)|^2 \delta(\Sigma p_i - \Sigma p_f) \prod_f \frac{d^3p_f}{(2\pi)^3}.$$

Integrating this expression over one of the momenta $\underline{p}_i$, we eliminate three of the three-dimensional δ -functions and obtain

$$d\omega = \frac{1}{(2\pi)^4} |M(p_i, p_f)|^2 \delta(\Sigma \epsilon_i - \Sigma \epsilon_f) \prod_f \frac{d^3p_f}{(2\pi)^3}, \quad (28.3)$$

where Π' is the product Π with one of the factors $\frac{d^3p_f}{(2\pi)^3}$ missing, and the ϵ_i and ϵ_f are the particle energies in the initial and final states.

Writing any one of the factors in Π' in the form

$$\frac{d^3p_f}{(2\pi)^3} = p_f d\epsilon_f d\Omega_f,$$

where $d\Omega_f$ is the element of solid angle about $\underline{p}_f$, and integrating over $\frac{d\epsilon_f}{\epsilon_f}$, we eliminate the energy δ -function, and obtain

$$d\omega = \frac{1}{(2\pi)^4} |M(p_i, p_f)|^2 p_f d\Omega_f \prod_f \frac{d^3p_f}{(2\pi)^3}, \quad (28.4)$$

where Π' is the same as Π except that the above factor is missing, and $p_f = \int p_f \delta(\Sigma \epsilon_i - \Sigma \epsilon_f) d\epsilon_f$ (in general p_f differs from p_e).

Dividing $d\omega$ by the incident current density J , we obtain the differential scattering cross section $d\sigma$,

$$d\sigma = \frac{d\omega}{J}. \quad (28.5)$$

2. Summation and Averaging Over Photon and Electron Polarizations.

When we are not interested in polarization eigenstates of the scattered particles (i.e., in the electron spins and photon polarizations), expression (28.4) should be summed over all possible polarizations in the final state. If, in addition, the particles are not polarized in the initial state, (28.4) should be averaged over all polarizations in the initial state.

We shall show first how to sum and average over electron spins. For simplicity, let us consider a case in which there is only one electron (or positron) both in the initial and in the final state. In this case the quantity $M(\underline{p}_i, \underline{p}_f)$ which enters the matrix element $S_{fi}^{(n)}$ can be written

$$M(\underline{p}_i, p_f) = \bar{u}_2 Q u_1,$$

where u_1 and u_2 are the spinor amplitudes of the electron in the initial and final states, respectively, and Q is some matrix. We wish to find

$$\sum_s |M(p_i, p_f)|^2 = \frac{1}{2} \sum_s \sum_s (u_2^\dagger Q u_1) (u_1^\dagger Q^\dagger u_2),$$

where $\sum_s$ and $\sum_s$ denote summation over the two electron spins in the final and initial states. According to (9.20),

$$\sum_s |M(p_i, p_f)|^2 = \frac{1}{8\pi^2} \text{Sp} \{ Q (\hat{p}_1 - m) \bar{Q} (\hat{p}_2 - m) \}, \quad (28.6)$$

where $\mathbf{p}_1$ and $\mathbf{p}_2$ are the momenta of the electron in the initial and final states, ϵ_1 and ϵ_2 are the corresponding energies, and

$$\bar{Q} = \beta Q \beta. \quad (28.7)$$

We shall now show how to sum over various photon polarizations. For simplicity let us consider a case in which only one photon takes part, and let its polarization be given by the four-vector $\hat{\epsilon}$. In this case, Q can clearly be written

$$Q = \hat{\epsilon} G,$$

where G is some matrix independent of the photon polarization. Noting that $\hat{\epsilon} = -\hat{\epsilon}$ (here $\hat{\epsilon}$ is a transverse vector with $\epsilon_4 = 0$), so that

$$\bar{Q} = -\bar{G} \hat{\epsilon},$$

we obtain the following expression for $\sum_{\epsilon} |M(p_i, p_f)|^2$:

$$\sum_{\epsilon} |M(p_i, p_f)|^2 = -\frac{1}{8\epsilon_1 \epsilon_2} \text{Sp} \{ \hat{\epsilon} G (\hat{p}_1 - m) \bar{G} \hat{\epsilon} (\hat{p}_2 - m) \}. \quad (28.8)$$

If we are not interested in the photon polarization, (28.8) should be summed or averaged over two independent directions of $\hat{\epsilon}$. Let us use $\hat{k}$ to denote the space-part of the photon wave vector and choose our coordinate system so that the $\hat{x}$ axis lies along $\hat{k}$. Then the two independent polarization vectors can be chosen as $\hat{\epsilon}^1(1, 0, 0)$ and $\hat{\epsilon}^2(0, 1, 0)$, and the desired sum can be written

$$\begin{aligned} \sum_{\epsilon} |M(p_i, p_f)|^2 &= -\frac{1}{8\epsilon_1 \epsilon_2} \sum_{r=1}^2 \text{Sp} \{ \hat{\epsilon}^r G (\hat{p}_1 - m) \bar{G} \hat{\epsilon}^r (\hat{p}_2 - m) \} = \\ &= -\frac{1}{8\epsilon_1 \epsilon_2} \sum_{r=1}^2 \text{Sp} \{ \gamma_r G (\hat{p}_1 - m) \bar{G} \gamma_r (\hat{p}_2 - m) \} \end{aligned} \quad (28.9)$$

($\sum_{\epsilon}$ denotes summation over the photon polarizations).

We shall show that the summation in this expression can be taken from $r=1$ to $r=4$, i.e., that

$$\sum_{\epsilon} |M(p_i, p_f)|^2 = -\frac{1}{8\epsilon_1 \epsilon_2} \text{Sp} \{ \gamma_4 G (\hat{p}_1 - m) \bar{G} \gamma_4 (\hat{p}_2 - m) \}. \quad (28.10)$$

To do this, we note that the equations of quantum electrodynamics are gauge invariant. Therefore, if the four-dimensional photon polarization $\hat{\epsilon}$ is replaced by $\hat{k}$ in a matrix element, we obtain zero. If, therefore, we replace one of the vectors $\hat{\epsilon}$ by $\hat{k}$ in (28.8), which is quadratic with respect to the matrix element, this expression vanishes and we get

$$\begin{aligned} \text{Sp} \{ (\gamma_4 + \gamma_1) G (\hat{p}_1 - m) \bar{G} \hat{\epsilon} (\hat{p}_2 - m) \} &= 0, \\ \text{Sp} \{ \hat{\epsilon} G (\hat{p}_1 - m) \bar{G} (\gamma_4 + \gamma_1) (\hat{p}_2 - m) \} &= 0, \end{aligned}$$

where the remaining vector $\hat{\epsilon}$ can be chosen arbitrarily. Choosing for $\hat{\epsilon}$ first γ_2 , and then γ_3 , we obtain

$$\begin{aligned} \text{Sp} \{ (\gamma_4 + \gamma_1) G (\hat{p}_1 - m) \bar{G} \gamma_2 (\hat{p}_2 - m) \} &= 0, \\ \text{Sp} \{ \gamma_4 G (\hat{p}_1 - m) \bar{G} (\gamma_4 + \gamma_1) (\hat{p}_2 - m) \} &= 0, \end{aligned}$$

from which it follows that

$$\text{Sp} \{ \gamma_2 G (\hat{p}_1 - m) \bar{G} \gamma_2 (\hat{p}_2 - m) \} + \text{Sp} \{ \gamma_3 G (\hat{p}_1 - m) \bar{G} \gamma_3 (\hat{p}_2 - m) \} = 0.$$

Adding this expression multiplied by $-\frac{1}{8\epsilon_1 \epsilon_2}$ to (28.9), we obtain (28.10).

This is in agreement with the general results of Section 15, where it was shown that "longitudinal" and "scalar" photons are not emitted.

A similar expression holds for a process involving several photons. If, for instance, two photons with polarizations $\hat{\epsilon}_1$ and $\hat{\epsilon}_2$ participate, the matrices Q and $\bar{Q}$ can be written

$$Q = \hat{\epsilon}_1 \hat{\epsilon}_2 + \hat{\epsilon}_1 \hat{\epsilon}_3, \quad \bar{Q} = \hat{\epsilon}_1 \bar{G} \hat{\epsilon}_2 + \hat{\epsilon}_2 \bar{G} \hat{\epsilon}_1.$$

Repeating the above considerations, it is easily shown that by summing $\sum_{\epsilon} |M(p_i, p_f)|^2$ over the various photon polarizations we obtain

$$\begin{aligned} \sum_{\epsilon} |M(p_i, p_f)|^2 &= -\frac{1}{8\epsilon_1 \epsilon_2} \sum_{r=1}^2 \text{Sp} \{ (\hat{\epsilon}_1^r G \hat{\epsilon}_2^r + \hat{\epsilon}_1^r \hat{\epsilon}_3^r) (\hat{p}_1 - m) (\hat{\epsilon}_1^r \bar{G} \hat{\epsilon}_2^r + \hat{\epsilon}_1^r \bar{G} \hat{\epsilon}_3^r) \times \\ &\times (\hat{p}_2 - m) \} = -\frac{1}{8\epsilon_1 \epsilon_2} \text{Sp} \{ (\gamma_1 G \gamma_1 + \gamma_1 G \gamma_3) (\hat{p}_1 - m) (\gamma_1 \bar{G} \gamma_1 + \gamma_1 \bar{G} \gamma_3) (\hat{p}_2 - m) \}. \end{aligned} \quad (28.11)$$

Summing over the four values of μ and ν leads to great simplification in calculation of $\sum_{\epsilon} |M(p_i, p_f)|^2$. Indeed, using the relations [see (9.24)]

$$\left. \begin{aligned} \gamma_a \gamma_b &= 4s, \\ \gamma_a \hat{a} \gamma_b &= -2\hat{a}, \\ \gamma_a \hat{a} \hat{b} \gamma_b &= 4ab, \\ \gamma_a \hat{a} \hat{b} \hat{c} \gamma_b &= -2\hat{c}ab, \end{aligned} \right\} \quad (28.12)$$

where $\hat{a}$ is an arbitrary scalar and a, b, c are arbitrary vectors, makes it possible to decrease the number of matrices involved in the trace in the expression for $\sum_{\mathbf{p}} |\mathbf{M}|^2$ (their number is reduced by $2p$, where p is the number of photons). Traces can be calculated with the aid of the relations [see (9.23)].

$$\frac{1}{4} \text{Sp } \hat{A}_1 \hat{A}_2 = A_1 A_2, \quad (28.13)$$

$$\frac{1}{4} \text{Sp } \hat{A}_1 \hat{A}_2 \hat{A}_3 \hat{A}_4 = (A_1 A_2)(A_3 A_4) + (A_1 A_4)(A_2 A_3) - (A_1 A_3)(A_2 A_4),$$

where $\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{A}_4$ are arbitrary four-vectors. Clearly the trace of an odd number of factors $\hat{A}_i$ vanishes. We make note also of the following useful formula:

$$\begin{aligned} \frac{1}{4} \text{Sp } (\hat{A}_1 + a_1)(\hat{A}_2 + a_2)(\hat{A}_3 + a_3)(\hat{A}_4 + a_4) = \\ = (A_1 A_2 + a_1 a_2)(A_3 A_4 + a_3 a_4) + (A_1 A_4 + a_1 a_4)(A_2 A_3 + a_2 a_3) - \\ - (A_1 A_3 + a_1 a_3)(A_2 A_4 + a_2 a_4), \end{aligned} \quad (28.13')$$

where the a_i are numbers.

3. Probabilities in the Presence of an External Field.

We shall show also how to determine the probabilities for transitions in the presence of an external electromagnetic field.

If one electron takes part in the process, in the first approximation the matrix element can be written

$$S_{i \rightarrow f}^{(1)} = e \bar{u}_2 \left(\int_{-\infty}^{\infty} e^{-iqx} \hat{A}^{(e)}(x) d^4 x \right) u_1, \quad (28.14)$$

where $q = p_2 - p_1$. Let us first assume that the external field is time-independent. Then

$$S_{i \rightarrow f}^{(1)} = e (\bar{u}_2 \hat{a} u_1) 2\pi \delta(\epsilon_2 - \epsilon_1),$$

where

$$\hat{a} = \int_{-\infty}^{\infty} \hat{A}^{(e)}(x) e^{-iqx} d^4 x$$

and ϵ_1 and ϵ_2 are the electron energies in the initial and final states. The transition probability is $|\hat{a}|^2$. Proceeding as in deriving (28.2), we obtain the following expression for the transition probability per unit time:

$$w_1 = 2\pi |\bar{u}_2 \hat{a} u_1|^2 \delta(\epsilon_2 - \epsilon_1). \quad (28.15)$$

In particular, if $\hat{A} = 0$, $\hat{A}_4 = \hat{1}$, we have

$$w_1 = 2\pi |\bar{u}_2 \hat{a} u_1|^2 \delta(\epsilon_2 - \epsilon_1). \quad (28.16)$$

We have obtained the well-known quantum mechanical expression for the transition probability in a constant field.

Let us also consider the case in which the external field is a periodic function of time. Inserting $\hat{A}^{(e)}(x) = \hat{e} e^{-i\omega t} \hat{A}^{(e)}(\mathbf{x})$ into (28.14) and eliminating one of the δ -functions as in (28.2), we obtain

$$w_1 = 2\pi |\bar{u}_2 \hat{a} u_1|^2 \delta(\epsilon_2 - \epsilon_1 - \omega).$$

4. Effective Perturbation Energy.

In the previous paragraphs we have stated that an external field cannot always be treated as a perturbation, and that in such cases the exact electron wave functions in the external field should be used in calculating the S-matrix elements (the external field is assumed constant). The matrix elements will then not contain three-dimensional δ -functions which give the law of conservation of momentum, but they will still contain a δ -function of the difference between ϵ_f and ϵ_i , which expresses the law of conservation of energy. We shall write the matrix element in the form

$$S_{i \rightarrow f}^{(1)} = -2\pi i \delta(\epsilon_f - \epsilon_i) U_{i \rightarrow f}. \quad (28.17)$$

The probability per unit time for the transition $i \rightarrow f$ is related to $U_{i \rightarrow f}$ by the expression

¹⁾ [The subscript q in the next three equations should be boldface. - editor's note.]

$$w_1 = 2\pi |U_{i \rightarrow f}|^2 \delta(\epsilon_i - \epsilon_f). \quad (28.18)$$

We note that $S_{i \rightarrow f}^{(n)}$ can be treated as the limit of

$$S_{i \rightarrow f}^{(n)}(t) = -i U_{i \rightarrow f} \int_{-\infty}^t e^{i\epsilon_f t'} dt' \quad \forall \epsilon_f - \epsilon_i$$

as $t \rightarrow \infty$. This function satisfies the well-known quantum mechanical equation of first-order perturbation theory, namely

$$i \frac{\partial S_{i \rightarrow f}^{(n)}(t)}{\partial t} = U_{i \rightarrow f} e^{i\epsilon_f t}.$$

Thus, U may be called the effective perturbation energy.

UNCLASSIFIED

STAT

AEC-tr-2876 (Pt. II)

UNITED STATES ATOMIC ENERGY COMMISSION

QUANTUM ELECTRODYNAMICS

(Part II)

A. I. Akhiezer and V. B. Berestetsky

Translated by:

CONSULTANTS BUREAU, INC.

Technical Information Service Extension, Oak Ridge, Tenn.



UNCLASSIFIED

RECEIVED
U. S. ATOMIC ENERGY COMMISSION
DIVISION OF INTELLIGENCE

JUN 4 1958

A.M. 7 8 9 10 11 12 P.M. 1 2 3 4 5 6

A translation of:

Kvantovaya Elektrodinamika, A. I. Akhiezer and V. B. Berestetsky Moskva, Gosudarstvennoe Izdatel'stvo Tekhniko-Teoreticheskoi Literatury, 1963, 428 pp.

Printed in the U.S.A. Price \$ 2.65. Bound in two parts to be sold together. Available from the Office of Technical Services, Department of Commerce, Washington 25, D.C.

CHAPTER VI

THE INTERACTION OF ELECTRONS AND PHOTONS

§ 29. Scattering of a Photon by an Electron.

1. Impossibility of First-Order Processes for the Free Electron.

We shall now go on to a systematic investigation of various phenomena which take place as a result of the interaction between electrons and the electromagnetic field. In each case the calculations will be made in the first nonvanishing approximation. Corrections due to higher approximations will be considered in Chapter VIII.

In the first order the S matrix can be written

$$S^{(1)} = -e \int \bar{\psi} \hat{A} \psi d^4x \quad (\hat{A} = \gamma_\mu A_\mu).$$

This expression contains the electromagnetic field operator $\hat{A}$ linearly. This means that the matrix elements of $S^{(1)}$ describe processes of emission or absorption of a photon¹⁾ such as, for instance, capture of a photon by an electron in a collision process. It is, however, not difficult to show that the matrix elements of $S^{(1)}$ vanish if the electron involved in the process is free.

Indeed, if the electron and photon are in eigenstates of their momenta, the matrix element for photon capture involves the integral

$$\int e^{i(p_1 + k - p_2)x} d^4x = (2\pi)^4 \delta(p_1 + k - p_2),$$

where p_1 , p_2 , and k are the four-momenta, respectively, of the electron in the initial state, of the electron in the final state, and of the absorbed photon. Thus the matrix element is different from zero only when momentum and energy are conserved according to

$$p_1 + k = p_2. \quad (29.1)$$

These conditions are incompatible, however, as we shall now show. Let us square both sides of (29.1),

$$p_1^2 + k^2 + 2p_1 k = p_2^2$$

¹⁾ We shall not here consider processes not involving photons, in which $\hat{A}$ is an external electromagnetic field. In this case the matrix elements of $S^{(1)}$ give, for instance, electron scattering in the Born approximation (see Section 13).

and make use of the fact that

$$p_1^2 = p_2^2 = -m^2, \\ k^2 = 0.$$

This gives

$$p_1 k = 0,$$

which cannot be satisfied, since

$$p_1 k < \sqrt{m^2 + |p_1|^2} |k|.$$

In order for the conservation laws to be maintained, a third body must take part in the interaction. Photon capture can take place only as a result of "triple collision" in which a significant role is played by the electron's interaction with another body. In a similar way it is not difficult to show that a photon cannot be emitted by a free electron, and electron-positron pair creation or annihilation by a single photon is impossible.

We are led to the conclusions that the simplest interaction processes are described by the second-order S matrix, which according to (20.15) can be written

$$S^{(2)} = \frac{e^2}{2} \int P [\bar{\psi}(1) \hat{A}(1) \psi(1), \bar{\psi}(2) \hat{A}(2) \psi(2)] d^4 x_1 d^4 x_2.$$

Since $S^{(2)}$ contains $\hat{A}$ bilinearly, its elements describe processes in which the total number of photons (before and after collision) is either two or zero. The latter case involves the interaction of electrons without photons, which shall be investigated in Chapter VII.

The matrix elements of $S^{(2)}$ involving two photon states can be divided into three types according to the total number of electrons in the initial and final states. This number can be either four, two, or zero. Of greatest physical interest are those matrix elements involving two electron states. Matrix elements with four electron states are simply products of first order $S^{(1)}$ matrix elements (see Section 23, Fig. 10, diagram 1). Matrix elements involving no electron states define the "self-mass" of the photon (see Section 23, Fig. 10, diagram 4).

Let us use $S^{(2)}$ to denote that part of $S^{(2)}$ which refers to processes involving two photon and two electron states. According to the results of Section 22 [see (22.30)],

$$iS^{(2)} = \frac{e^2}{2} \int \bar{\psi}(2) \hat{A}(2) S^F(21) \hat{A}(1) \psi(1) d^4 x_1 d^4 x_2, \quad (29.2)$$

where S^F is given by Eqs. (18.31) and (18.33). The simplest process described by S^F is the scattering of a photon by an electron, which we shall now consider in detail.

2. Photon Scattering by a Free Electron.

Let the wave functions of the initial and final electron states be ψ_1 and ψ_2 , respectively, and the potentials of the initial and final photon states be A_1 and A_2 . Then the matrix element of (29.2) which corresponds to this

transition is obtained by replacing $\bar{\psi}(2)$ by $\bar{\psi}_2(2)$, $\psi(1)$ by $\psi_1(1)$ and $A(2)$ by $A_2(2)$, $A(1)$ by $A_1(1)$ [see also Equations (23.2) and (24.9)], so that

$$iS_{12 \rightarrow r}^{(2)} = \frac{e^2}{2} \int \bar{\psi}_2(2) [\hat{A}_2(2) S^F(21) \hat{A}_1(1) + \hat{A}_1(2) S^F(21) \hat{A}_2(1)] \psi_1(1) d^4 x_2 d^4 x_1. \quad (29.3)$$

We shall choose ψ_1 and A_1 ($i = 1, 2$) to correspond to momentum and polarization eigenstates normalized for a unit volume:

$$\psi_i = u_i e^{ip_i x}, \quad A_i = \frac{e_i}{\sqrt{2\omega_i}} e^{ik_i x}, \quad (29.4)$$

where p_i and k_i are the electron and photon four momenta, $\omega_i = |k_i|$ is the frequency of the photon, and u_i and e_i are unit (spinor and vector) amplitudes.

Inserting (29.4) and (18.33) into (29.3), we obtain the following result, which is in agreement with the general rules formulated in Section 24 [compare Equation (24.9)],

$$S_{12 \rightarrow r}^{(2)} = \frac{-ie^2}{2\sqrt{\omega_1 \omega_2}} u_2 \left\{ \hat{e}_2 \frac{i\hat{f}_1 - m}{f_1^2 + m^2} \hat{e}_1 + \hat{e}_1 \frac{i\hat{f}_2 - m}{f_2^2 + m^2} \hat{e}_2 \right\} \times \\ \times u_1 (2\pi)^4 \delta(p_1 + k_1 - p_2 - k_2), \quad (29.5)$$

where

$$f_1 = p_1 + k_1 = p_2 + k_2, \quad f_2 = p_1 - k_2 = p_2 - k_1. \quad (29.6)$$

Figure 32 shows two diagrams corresponding to the two terms of the matrix element (29.3), (29.5); these are constructed according to the rules of Section 23. The lines on each of the diagrams are ordered according to the way the amplitudes are ordered in (29.5). The time axis is directed from right to left; in the first diagram the absorption takes place "first" and the emission "second"; on the second diagram this order is reversed.

As can be seen from (29.6), the matrix element differs from zero only when energy and momentum are conserved, that is when

$$p_1 + k_1 = p_2 + k_2. \quad (29.7)$$

As is always the case for collision of two particles, the conservation laws make it possible to determine four of the six components of the momenta in the final state from the initial values p_1 and k_1 ; for instance, we know p_2 and ω_2 for a given direction of k_2 .

Let us find the frequency ω_2 of the scattered photon as a function of its direction. We do this by squaring (29.7), obtaining

$$p_1^2 + k_1^2 + 2p_1 k_1 = p_2^2 + k_2^2 + 2p_2 k_2.$$

Since

$$p_1^2 = p_2^2 = -m^2, \quad k_1^2 = k_2^2 = 0,$$

we have

$$p_1 k_1 = p_2 k_2,$$

or eliminating p_2 according to (29.7), and obtaining

$$p_1 k_1 = p_1 k_2 + k_1 k_2, \quad (29.8)$$

we arrive at

$$\omega_1 (1 - v_1 \cos \theta_1) = \omega_2 (1 - v_1 \cos \theta_2) + \frac{\omega_1 \omega_2}{\epsilon_1} (1 - \cos \theta), \quad (29.9)$$

where

$$v_1 = \frac{|p_1|}{\epsilon_1}$$

is the initial velocity of the electron, ϵ_1 is its initial energy, θ is the angle between the initial and scattered photon,

$$\cos \theta = \frac{k_1 k_2}{\omega_1 \omega_2},$$

and θ_1 and θ_2 are the angles between the direction of the initial electron and those of the incident and scattered photons, respectively,

$$\cos \theta_1 = \frac{p_1 k_1}{|p_1| \omega_1}, \quad \cos \theta_2 = \frac{p_1 k_2}{|p_1| \omega_2}.$$

For the case of scattering on a stationary electron ($v_1 = 0, \epsilon_1 = m$) Equation (29.9) reduces to the well-known Compton scattering formula.

$$\omega_2 = \frac{\omega_1}{1 + \frac{\omega_1}{m} (1 - \cos \theta)}, \quad (29.10)$$

The quantities f_1 and f_2 are the four-momenta of the virtual electrons, for which the usual relation between the momentum and energy does not hold, so that

$$f_s^2 + m^2 \neq 0 \quad s = 1, 2.$$

The deviation of the virtual electron from a real one can be characterized by the coefficient

$$\kappa_s = \frac{f_s^2 + m^2}{m^2} \quad s = 1, 2,$$

which vanishes for a real electron [we note that the quantities $m^2 \kappa_s$ are the denominators in (29.5)]. From (29.6) it follows that

$$\left. \begin{aligned} m^2 \kappa_1 &= 2p_1 k_1 = 2p_2 k_1, \\ m^2 \kappa_2 &= -2p_1 k_2 = -2p_2 k_2. \end{aligned} \right\} \quad (29.11)$$

If before the collision the electron was at rest, then

$$\kappa_1 = -\frac{2\omega_1}{m}, \quad \kappa_2 = \frac{2\omega_2}{m}. \quad (29.12)$$

Let us find the cross section for this process. According to the general rules of Section 28, the differential scattering cross section for the final momenta in the intervals $d\mathbf{p}_2$ and $d\mathbf{k}_2$ is given by

$$d\sigma = \frac{e^4}{4\omega_1 \omega_2} |\bar{u}_2 Q u_1|^2 \frac{d\mathbf{p}_2 d\mathbf{k}_2}{f(2\pi)^3} \delta(p_1 + k_1 - p_2 - k_2) \delta(\epsilon_1 + \omega_1 - \epsilon_2 - \omega_2), \quad (29.13)$$

where Q is the operator enclosed by braces in (29.5), namely

$$Q = \frac{1}{m^2 \kappa_1} \hat{\epsilon}_2 (\hat{f}_1 - m) \hat{\epsilon}_1 + \frac{1}{m^2 \kappa_2} \hat{\epsilon}_1 (\hat{f}_2 - m) \hat{\epsilon}_2,$$

and $\mathbf{j}$ is the current density of the colliding particles. In the general case of scattering by a moving scattering center, the current density should be taken as

$$\mathbf{j} = p_1 (v_1 - v_1 \cos \theta),$$

In which ρ_1 and $\mathbf{v}_1$ are the density and velocity of the scattered particles, $\mathbf{v}_{11}$ is the velocity of the scatterer, and θ is the angle between $\mathbf{v}_1$ and $\mathbf{v}_{11}$. In our case

$$J = 1 - v_1 \cos \theta = -\frac{\rho_1 k_1}{\epsilon_1 \omega_1} = -\frac{m^2 \epsilon_1}{2 \epsilon_1 \omega_1}. \quad (29.14)$$

One of the δ -functions in (29.13) can be eliminated by performing the substitution

$$\delta(\rho_1 + k_1 - \rho_2 - k_2) d\rho_2 \rightarrow 1$$

and

$$\delta(\epsilon_1 + \omega_1 - \epsilon_2 - \omega_2) dk_2 \rightarrow \omega_2 d\omega_2 \left| \frac{\partial}{\partial \omega_2} (\epsilon_2 + \omega_2) \right|.$$

In the last expression ϵ_2 is a function of ω_2 since the momentum $\mathbf{p}_2$ is related to $\mathbf{k}_2$ by conservation of momentum:

$$\epsilon_2 = \sqrt{m^2 + (\mathbf{p}_1 + \mathbf{k}_1 - \mathbf{k}_2)^2}.$$

By differentiating (29.7), (29.9), and (29.11), we find that

$$\left| \frac{\partial (\epsilon_2 + \omega_2)}{\partial \omega_2} \right| = \frac{m^2 |\epsilon_1|}{2 \omega_2 \epsilon_2}.$$

Thus (29.13) can be rewritten

$$d\sigma = 4\pi^2 \frac{\omega_2^3}{n^2} \frac{\epsilon_1 \epsilon_2}{x_1^2} |\bar{u}_2 Q u_1|^2 d\omega_2, \quad (29.15)$$

where

$$r_0 = \frac{e^2}{4\pi m}$$

is the "classical electron radius".¹⁾

If the initial electron state is not a spin eigenstate and if we are not concerned with the electron polarization in the final state, (29.15) should be summed over the final and averaged over the initial electron spins. The quantity thus obtained, namely

¹⁾ We note that $\frac{e^2}{\sqrt{4\pi}}$ is the electron charge in Gaussian units.



Fig. 32

$$\frac{1}{2} \sum_{\mu_1, \mu_2} d\sigma$$

(μ_1 and μ_2 denote the electron spin orientations in the initial and final states), shall be denoted briefly $d\sigma$. According to the general equation (9.20).

$$d\sigma = \frac{r_0^2}{2} \frac{\omega_2^3}{m^2} \frac{1}{x_1^2} \text{Sp} [(\hat{p}_1 - m) \bar{Q} (\hat{p}_2 - m)] d\omega_2 \quad (\bar{Q} = \beta Q + \beta), \quad (29.16)$$

Inserting Q into this expression and making use of the fact that any four-vector q , with three real components and one (the fourth) imaginary, satisfies

$$\bar{q} = -\hat{q}, \quad (29.17)$$

we arrive at

$$\begin{aligned} Q(\hat{p}_1 - m) \bar{Q}(\hat{p}_2 - m) = \\ = \left[\frac{1}{m^2 \epsilon_1} \hat{\epsilon}_2 (\hat{p}_1 - m) \hat{\epsilon}_1 + \frac{1}{m^2 \epsilon_2} \hat{\epsilon}_1 (\hat{p}_2 - m) \hat{\epsilon}_2 \right] (\hat{p}_1 - m) \times \\ \times \left[\frac{1}{m^2 \epsilon_1} \hat{\epsilon}_1 (\hat{p}_1 - m) \hat{\epsilon}_2 + \frac{1}{m^2 \epsilon_2} \hat{\epsilon}_2 (\hat{p}_2 - m) \hat{\epsilon}_1 \right] (\hat{p}_2 - m). \end{aligned} \quad (29.18)$$

If the initial photon is also unpolarized, we obtain the scattering cross section irrespective of the final photon polarization by summing (29.16) over the final, and averaging it over the initial, photon polarization states. The quantities so obtained, namely

$$\frac{1}{2} \sum_{\nu_1, \nu_2} d\sigma$$

(ν_1 and ν_2 are the polarization indices of the photon), we shall denote briefly by $d\sigma$.

The calculation of this cross section becomes quite simple, since the summation can be taken over four, rather than two, photon polarizations, including the "longitudinal" and "scalar" polarizations (see Section 28). Then we need only perform the substitution

$$\hat{\epsilon}_1 = \gamma_4, \quad \hat{\epsilon}_2 = \gamma_j \quad (29.19)$$

in (29.18), and sum over j and j (where $j, \bar{j} = 1, 2, 3, 4$).

Thus we obtain

$$d\sigma = r_0^2 \frac{\omega_2^2}{m^2 \omega_1^2} d\omega_2 \frac{1}{\omega_1} \text{Sp } F, \quad (29.20)$$

where F is the operator given by (29.18) with the substitution (29.19).¹⁾

The calculation of $\text{Sp } F$ is not difficult. In the first place, Equations (7.27) and (28.12) can be used to eliminate γ_i in products such as $\gamma_1 \cdots \gamma_i$. After this is done, each term of F will contain no more than four γ -matrices. The traces of such products have been calculated in Section 9 [see (9.23)]. The result can then be written

$$\frac{1}{8} \text{Sp } F = 4 \left(\frac{1}{\omega_1} + \frac{1}{\omega_2} \right)^2 - 4 \left(\frac{1}{\omega_1} + \frac{1}{\omega_2} \right) - \left(\frac{\omega_1}{\omega_2} + \frac{\omega_2}{\omega_1} \right). \quad (29.21)$$

3. Angular Distribution and Total Cross Section for Unpolarized Photons.

According to (29.20), (29.21), (29.10), and (29.12), the differential cross section for scattering of a photon by an electron at rest can be written²⁾

$$d\sigma = \frac{r_0^2}{2} \left(\frac{\omega_2}{\omega_1} \right)^2 \left(\frac{\omega_1}{\omega_2} + \frac{\omega_2}{\omega_1} - \sin^2 \theta \right) d\omega_2. \quad (29.22)$$

In this formula, ω_2 is a function of ω_1 and θ . Using (29.10), we obtain

$$d\sigma = \frac{r_0^2}{2} (1 + \cos^2 \theta) \frac{1}{\left[1 + \frac{\omega_1}{m} (1 - \cos \theta) \right]^2} \left[1 + \frac{\left(\frac{\omega_1}{m} \right)^2 (1 - \cos \theta)^2}{(1 + \cos^2 \theta) \left[1 + \frac{\omega_1}{m} (1 - \cos \theta) \right]} \right] d\omega_2. \quad (29.23)$$

Figure 33 gives $\frac{1}{\omega_2} \frac{d\sigma}{d\omega_2} \frac{\omega_2}{\omega_1}$ (the relative intensity) as a function of θ for various values of $\omega_1/m \equiv w$.

For photon energies $\omega_1 \ll m$, (29.23) goes over into the classical Thomson scattering formula

$$d\sigma = \frac{r_0^2}{2} (1 + \cos^2 \theta) d\omega_2, \quad \omega_1 \ll m. \quad (29.24)$$

If the photon energy is much larger than the electron rest energy, so that $\omega_1 \gg m$, (29.23) can be used to obtain simple expressions for two limiting cases. When $\theta \ll \sqrt{\frac{m}{\omega_1}}$,

¹⁾ We note that (29.20) shows that $d\sigma$ is an invariant. Indeed, $\text{Sp } F$ and κ_1 are invariants. The quantity $\omega^2 d\omega$ is also an invariant since it transforms like

$$\omega d\omega d\omega' \rightarrow \frac{d^3k}{m} = 2\pi (k^2)^2 dk.$$

²⁾ O. Klein and Nishina, Z. Physik 82,853 (1929); I. Tamm, Z. Physik 62,545 (1930).

$$d\sigma = r_0^2 d\omega_2, \quad \omega_1 \gg m, \quad \theta \ll \sqrt{\frac{m}{\omega_1}}. \quad (29.25)$$

When $\theta \gg \sqrt{\frac{m}{\omega_1}}$

$$d\sigma = \frac{r_0^2}{2} \frac{m}{\omega_1} \frac{d\omega_2}{1 - \cos \theta}, \quad \omega_1 \gg m, \quad \theta \gg \sqrt{\frac{m}{\omega_1}}. \quad (29.26)$$

We note that when $\theta = \sqrt{\frac{m}{\omega_1}}$ (with $\frac{m}{\omega_1} \ll 1$) Equations (29.25) and (29.26) give the same result. This shows that (29.25) is a good approximation for the angular distribution of (29.23) when $\omega_1 \gg m$ for angles in the interval $0 \leq \theta \leq \sqrt{\frac{m}{\omega_1}}$, and that (29.26) approximates this distribution when $\theta \gg \sqrt{\frac{m}{\omega_1}}$.

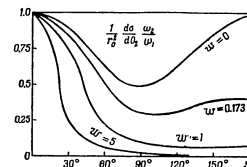


Fig. 33

Integration of (29.23) leads to the following expression for the total scattering cross section:

$$\sigma = 2\pi r_0^2 \left\{ \frac{1+w}{w^2} \left[\frac{2w(w+1)}{1+2w} - \ln(1+2w) \right] + \frac{1}{2w} \ln(1+2w) - \frac{1+3w}{(1+2w)^2} \right\}, \quad (29.27)$$

where

$$w = \omega_1/m.$$

For $\omega_1 \ll m$ (29.27) can be expanded in a power series in w , whence

$$\sigma = \frac{8\pi}{3} r_0^2 (1 - 2w + \dots), \quad \omega_1 \ll m, \quad (29.28)$$

and the cross section is approximately constant.

When $\omega_1 \gg m$,

$$\sigma = \pi r_0^2 \frac{m}{\omega_1} \left(\ln \frac{2\omega_1}{m} + \frac{1}{2} \right), \quad \omega_1 \gg m. \quad (29.29)$$

Thus the cross section varies essentially inversely as the photon energy. Figure 34 gives the dependence of $\sigma/\frac{8\pi}{3}r_0^2$ on ω according to (29.27).

Angular Distribution for Scattering of Polarized Photons.

Let us go on to the question of polarized photon scattering. In this case, calculation of the cross section according to (29.16) is somewhat more difficult than it is for unpolarized photons. However, the commutation rules for the γ matrices as given by (7.27) and (28.12) can be used here again to transform (29.18) into a form in which the factors $\hat{e}_1$ and $\hat{e}_2$ which occur twice in each term are always adjacent. We then use the fact that

$$\hat{e}_r \hat{e}_r = \hat{e}_r^2 = 1, \quad r = 1, 2$$

and each term, as previously, will contain products of no more than four γ matrices, so that we can again make use of (9.23).

We present here the final result for the differential cross section for scattering of polarized photons by an electron at rest:

$$d\sigma = \frac{r_0^2}{4} \left(\frac{\omega_2}{\omega_1} \right)^2 \left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} - 2 + 4 \cos^2 \theta \right] d\omega_2, \quad (29.30)$$

where θ is the angle between the directions of polarization of the incident and scattered photons

$$\cos \theta = \hat{e}_1 \hat{e}_2.$$

Two mutually perpendicular directions $\hat{e}_2$, suitable for describing two linearly independent states of the scattered photon, can be chosen as follows; one of these is perpendicular to $\hat{e}_1$:

$$\cos \theta^{(1)} = \hat{e}_1^{(1)} \hat{e}_2 = 0,$$

and the second is in the plane passing through $\hat{k}_2$ and $\hat{e}_1$ so that

$$\cos \theta^{(2)} = 1 - \sin^2 \theta \cos^2 \varphi,$$

where φ is the angle between the plane defined by $\hat{k}_1, \hat{k}_2$ and that defined by $\hat{k}_2, \hat{e}_1$. We shall denote the differential cross sections corresponding to these two states by $d\sigma^{(1)}$ and $d\sigma^{(2)}$. Their sum, averaged over φ , gives the unpolarized photon scattering cross section (29.20).

When $\omega_1 \ll m$

$$d\sigma^{(1)} = 0, \quad d\sigma^{(2)} = r_0^2 (1 - \sin^2 \theta \cos^2 \varphi) d\omega_2. \quad (29.31)$$

When $\omega_1 \gg m$, we must differentiate between the large and small angle regions. For $\theta \ll \sqrt{\frac{m}{\omega_1}}$ we get a result identical with (29.31). For $\theta \gg \sqrt{\frac{m}{\omega_1}}$

$$d\sigma^{(1)} = d\sigma^{(2)} = \frac{1}{2} \quad (29.32)$$

where $d\sigma$ is given by (29.26). This last result means that for large energies and large scattering angles the photon's polarization depends on the polarization of the incident photon.

In conclusion, we note that the photon scattering cross section contains the mass m of the scattering center in the denominator. Therefore when $m \rightarrow \infty$ no scattering takes place.

This conclusion, however, is inexact since it considers only the second-order S matrix. The fourth-order S matrix already has elements corresponding to photon scattering by infinitely heavy charges. Coherent photon scattering by nuclei, which corresponds to these matrix elements, will be considered in Section 47.

§ 30. Emission and Absorption of a Photon.

1. General Expression for the Matrix Element.

Let us return to the question of photon emission. We have seen above that the free electron does not radiate, and that emission can occur only when it interacts with other bodies. In many important cases this interaction can be described with the aid of an external field in the electron Hamiltonian. The electron wave function operator can be expanded in a series of eigenfunctions of this Hamiltonian [see (17.6)], and the electron state itself will automatically take account of its interaction with other particles. In this case photon emission can be studied with the aid of the first-order S matrix whose elements between the states thus defined are in general different from zero.

Let us write out the general expression for the matrix element corresponding to radiation. Let ψ_1 be the initial electron wave function, ψ_2 be the final electron wave function, and A be the electromagnetic potential corresponding to a definite state of the emitted photon. Then

$$S_{1 \rightarrow 2}^{(1)} = -e \int \bar{\psi}_2 \hat{A} \psi_1 d^4x \sqrt{N+1}. \quad (30.1)$$

Here N is the number of photons in the given state before the electron radiates. We shall assume below that $N = 0$.

If we are interested in the emission of a photon in momentum and polarization eigenstates, then inserting the potential of (29.4) into (30.1), we obtain

$$S_{1 \rightarrow 2}^{(1)} = -\frac{e}{\sqrt{2\omega}} \int \bar{\psi}_2 \hat{e} e^{ikx} \psi_1 d^4x = -\frac{ie}{\sqrt{2\omega}} \int \bar{\psi}_2 \alpha e \psi_1 e^{ikr - i\omega t} dr dt. \quad (30.2)$$

2. Electric Multipole Radiation.

Let us consider radiation by an electron whose initial and final states are bound and which emits a photon whose wave length is long in comparison with the dimensions of the region in which the electron motion takes place. Then, as we shall see below, the emission probability is simply related to the electric or magnetic multipole moment of the electron. Therefore the radiation of such a system is called multiple radiation.

The results which we shall obtain can be applied also to emission by a proton, and therefore to finding emission probabilities for nuclei. The fact that the field in which a proton moves in the nucleus is not electromagnetic is of no significance. We shall use the nonrelativistic approximations for the wave functions.¹⁾

We shall start by using expression (30.1) for the matrix element. If ψ_1 and ψ_2 correspond to stationary states, they depend on the time through the factors $e^{-i\omega_1 t}$ and $e^{-i\omega_2 t}$. Similarly, $\hat{A}$ contains the time in the form $e^{-i\omega t}$. Thus integration over time in the matrix element is simple to perform. Let us introduce the matrix element $U_{i \rightarrow f}$ according to (28.17):

$$S_{i \rightarrow f} = -2\pi i U_{i \rightarrow f} \delta(\epsilon_1 - \epsilon_2 - \omega). \quad (30.3)$$

Then

$$U_{i \rightarrow f} = -ie \int \hat{\psi}_2 \hat{A}^* \psi_1 dr, \quad (30.4)$$

where ψ_1 , ψ_2 and $\hat{A}$ are functions only of $\underline{r}$ (without their time dependent factors).

Let $\hat{A}$ correspond to an angular momentum and parity eigenstate of the photon ²⁾

$$A = A_{LM\lambda}$$

[where L is the angular momentum of the photon, M is its projection, $(-1)^{L-\lambda+1}$ is the parity of the state, and $\lambda = 0$ or 1 ; when $\lambda = 0$ we say the state is a magnetic one, and when $\lambda = 1$, it is an electric one]. Let us first consider the emission of a photon in an electric state. We shall here make use of that normalization for the potentials which corresponds to a photon whose energy lies in the interval $d\omega$. For this normalization we need not multiply the probability by the number of states in this interval. According to (5.19),

$$\hat{A} = \sqrt{\frac{\omega d\omega}{2(2\pi)^3}} \{ -\alpha \alpha_{LM}^{(0)} + C(\Phi_{LM} - \alpha \alpha_{LM}^{(0)}) \}, \quad (30.5)$$

where

$$\left. \begin{aligned} \alpha_{LM}^{(0)} &= \sqrt{\frac{L}{2L+1}} g_{L+1}(\omega r) Y_{L, L+1, M} + \sqrt{\frac{L+1}{2L+1}} g_{L-1}(\omega r) Y_{L, L-1, M} \\ \alpha_{LM}^{(1)} &= \sqrt{\frac{L+1}{2L+1}} g_{L+1} Y_{L, L+1, M} - \sqrt{\frac{L}{2L+1}} g_{L-1} Y_{L, L-1, M} \\ \Phi_{LM} &= g_L Y_{LM} \end{aligned} \right\} \quad (30.6)$$

¹⁾ It is here necessary to take into account the anomalous moment of the proton. It is possible also to consider radiation by a neutron in this approximation, treating it as a particle without charge but with a magnetic moment.

²⁾ Here we have changed the angular momentum quantum number designation of the photon from that used in Sections 4 and 5 (instead of j we shall here use L).

and C is an arbitrary constant. [The g_L are defined in (4.29), and the spherical vectors $\underline{Y}_{L, L, M}$ in (4.11)].

We note that Φ_{LM} and $\alpha_{LM}^{(-1)}$ are related by

$$\alpha_{LM}^{(-1)} = -\frac{i}{\omega} \nabla \Phi_{LM}, \quad \alpha_{LM}^{(-1)*} = \frac{i}{\omega} \nabla \Phi_{LM}^*, \quad (30.7)$$

which are the Fourier transforms of (4.14).

Let us now make use of the assumption that the electron is moving in a region small compared to the wave length of the photon. In this case we may assume that those values of L are important in (30.4) for which

$$\omega r \ll 1.$$

We can then limit ourselves to the first term in the series expression for the Bessel function, that is we may set

$$\left. \begin{aligned} g_L(\omega r) &= 4\pi \frac{(\omega r)^L}{(2L+1)!}, \\ (2L+1)! &= 1 \cdot 3 \cdot 5 \dots (2L+1). \end{aligned} \right\} \quad (30.8)$$

In (30.6) we may neglect terms containing g_{L+1} since it is much smaller than g_{L-1} :

$$\left| \frac{g_{L+1}}{g_{L-1}} \right| \approx \frac{(\omega r)^2}{(2L+1)(2L+3)} \ll 1.$$

In this approximation $\alpha_{LM}^{(1)}$ can also be given in terms of Φ_{LM} , namely

$$\alpha_{LM}^{(1)} = -\sqrt{\frac{L+1}{L}} \alpha_{LM}^{(-1)} = \frac{i}{\omega} \sqrt{\frac{L+1}{L}} \nabla \Phi_{LM}. \quad (30.9)$$

Inserting (30.5) and (30.9) into (30.4), we obtain the following expression for the matrix element:

$$U_{i \rightarrow f} = -e \sqrt{\frac{\omega d\omega}{2(2\pi)^3}} \int \psi_2^* \left[i \sqrt{\frac{L+1}{L}} \frac{1}{\omega} \nabla \Phi_{LM}^* \alpha + C(\Phi_{LM}^* - \frac{i}{\omega} \alpha \nabla \Phi_{LM}^*) \right] \psi_1 dr. \quad (30.10)$$

This expression can be simplified. The integral containing $\nabla \Phi$ can be written

$$\int \psi_2^* (-i\alpha \nabla \Phi_{LM}^*) \psi_1 dr = \int \psi_2^* (\alpha p \Phi_{LM}^* - \Phi_{LM}^* \alpha p) \psi_1 dr,$$

where p is the momentum operator. Since the only term in the Hamiltonian

$$H = \alpha p + \beta m - ie \beta \hat{A}^{(0)}$$

(which includes the external field) which does not commute with Φ_{LM} is the operator αp , we may write

$$\alpha p \Phi_{LM}^* - \Phi_{LM}^* \alpha p = H \Phi_{LM}^* - \Phi_{LM}^* H.$$

Making use of the fact that H is Hermitian, and that ψ_1 and ψ_2 are its eigenfunctions given by

$$\begin{aligned} H \psi_1 &= \epsilon_1 \psi_1, \\ H \psi_2 &= \epsilon_2 \psi_2, \\ \epsilon_1 - \epsilon_2 &= \omega, \end{aligned}$$

we arrive at

$$\int \psi_2^* (-i\alpha \nabla \Phi_{LM}^*) \psi_1 dr = -\omega \int \psi_2^* \Phi_{LM}^* \psi_1 dr.$$

This expression shows that the terms containing $\nabla \Phi$ in (30.10) cancel, which could have been expected from gauge invariance, and since the matrix elements should not depend on an arbitrary constant. Thus

$$U_{i \rightarrow f} = -e \sqrt{\frac{\omega}{2(2\pi)^3}} \sqrt{\frac{L+1}{L}} \int \psi_2^* \Phi_{LM}^* \psi_1 dr. \quad (30.11)$$

Inserting (30.6) for Φ_{LM} and (30.8) into this expression, we obtain

$$U_{i \rightarrow f} = -e \sqrt{\frac{\omega}{2(2\pi)^3}} (-i)^L \sqrt{\frac{L+1}{L}} \frac{e^L}{(2L+1)!} \int \psi_2^* \psi_1 r^L Y_{LM}^* dr.$$

We note that up to terms in v^2 (where v is the velocity of the electron) the wave functions can be replaced by nonrelativistic functions.

Let us now define the electric multipole moment operator $Q_{LM}^{(e)}$:

$$Q_{LM}^{(e)} = \frac{e}{\sqrt{4\pi}} \sqrt{\frac{4\pi}{2L+1}} r^L Y_{LM}^*. \quad (30.12)$$

Here $\frac{e}{\sqrt{4\pi}}$ is the charge in "ordinary" (not Heaviside) units. The matrix element (30.11) can be expressed in terms of the matrix element of the electric multipole moment

$$(Q_{LM}^{(e)})_{21} = \int \psi_2^* Q_{LM}^{(e)} \psi_1 dr.$$

We shall drop the indices corresponding to the initial and final states in this expression, and shall henceforth denote the matrix element, in the same way as the operator, $Q_{LM}^{(e)}$. Then

$$U_{i \rightarrow f} = -(-i)^L \sqrt{\frac{\omega}{2(2\pi)^3}} \sqrt{\frac{L+1}{L}} \frac{1}{(2L+1)!} Q_{LM}^{(e)}. \quad (30.13)$$

The transition probability is given in terms of $U_{i \rightarrow f}$ according to (28.18) by

$$w = 2\pi |U_{i \rightarrow f}|^2 \delta(\epsilon_1 - \epsilon_2 - \omega).$$

Thus the probability for emitting a photon with angular momentum L , projection of the angular momentum in a given direction M , and parity $(-1)^L$ is given by

$$w_{LM}^{(e)} = \frac{2(L+1)}{L(2L+1) [(2L-1)!]^2} \omega^{2L+1} |Q_{LM}^{(e)}|^2. \quad (30.14)$$

1) When $M = 0$, (30.12) reduces to

$$Q_{L0}^{(e)} = e' r^L P_L(\cos \theta) \quad \left(e' = \frac{e}{\sqrt{4\pi}} \right),$$

where P_L is a Legendre polynomial. When $L = 1$, this gives

$$Q_{10}^{(e)} = e' z,$$

and when $L = 2$,

$$Q_{20}^{(e)} = e' \frac{1}{2} (3z^2 - r^2),$$

which correspond to the usual definitions of the dipole and quadrupole moments.

In particular, when $L = 1$, we obtain the equation for dipole radiation:

$$\omega_{LM}^{(1)} = \frac{4}{3} \omega^3 |Q_{LM}^{(1)}|^2.$$

3. Magnetic Multipole Radiation.

Let us go on to a consideration of photon emission in a state of the magnetic type. According to (5.21), in this case

$$\lambda = -i\beta \sqrt{\frac{\omega}{2(2\pi)^3}} \alpha_{LM}^{(0)}, \quad (30.15)$$

where

$$\alpha_{LM}^{(0)} = g_L(\omega r) Y_{LM}^{(0)}. \quad (30.16)$$

Making use of (30.8), we obtain

$$U_{i \rightarrow f} = e \sqrt{\frac{\omega}{2\pi}} \frac{(-i)^L \omega^L}{(2L+1)!} \int \psi_2^* \alpha Y_{LM}^{(0)*} r^L \psi_1 dr. \quad (30.17)$$

If the electron velocity is small, the wave functions can be written in terms of the nonrelativistic two-component functions φ (see Section 14). In this case we obtain

$$e \int \psi_2^* \alpha Y_{LM}^{(0)*} dr = \int \varphi_2^* \left[\frac{e}{2m} (p \alpha_{LM}^{(0)} + \alpha_{LM}^{(0)*} p - \text{curl} \alpha_{LM}^{(0)*}) \right] \varphi_1 dr \quad (30.18)$$

(where μ is the magnetic moment of the electron). This last expression can be transformed into a different form. First, since

$$\text{div} \alpha_{LM}^{(0)} = 0,$$

we may write

$$p \alpha_{LM}^{(0)} + \alpha_{LM}^{(0)*} p = 2 \alpha_{LM}^{(0)*} p.$$

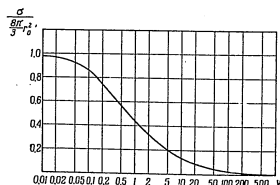


Fig. 34

Further we shall use the fact that ¹⁾

$$Y_{LM}^{(0)} = \frac{1}{\sqrt{L(L+1)}} [r \nabla Y_{LM}]$$

and therefore that

$$\alpha_{LM}^{(0)} = \frac{1}{\sqrt{L(L+1)}} [r \nabla \Phi_{LM}].$$

Thus

$$\alpha_{LM}^{(0)} p = \frac{1}{\sqrt{L(L+1)}} \nabla \Phi_{LM} L, \quad (30.19)$$

where $L = [p]$ is the orbital angular-momentum operator.

The last term in (30.18) can also be expressed in terms of Φ_{LM} . From (4.31) and (5.21) we have

$$\text{curl} \alpha_{LM}^{(0)} = i \omega \alpha_{LM}^{(1)}$$

or, making use of (30.9),

$$\text{curl} \alpha_{LM}^{(0)} = \sqrt{\frac{L+1}{L}} \nabla \Phi_{LM}. \quad (30.20)$$

Inserting (30.15) into (30.4), we obtain

$$U_{i \rightarrow f} = -\sqrt{\frac{\omega}{2\pi}} \int \varphi_2^* (\nabla \Phi_{LM}^*) \left[\frac{e}{\sqrt{L(L+1)}} L + \sqrt{\frac{L+1}{L}} p \right] \varphi_1 dr. \quad (30.21)$$

Let us define the magnetic multiple moment operator as follows:²⁾

$$Q_{LM}^{(0)} = \sqrt{\frac{4\pi}{2L+1}} \nabla (r^L Y_{LM}) \left(\frac{e'}{(L+1)m} L + \mu' \right) \quad (30.22)$$

$$\left(e' = \frac{e}{\sqrt{4\pi}}; \mu' = \frac{\mu}{\sqrt{4\pi}} \right).$$

¹⁾ This relation is easy to obtain from the recursion relation which gives the derivatives of the spherical function of order L in terms of those of order $L \pm 1$.

²⁾ When $M = 0$,

$$Q_{LM}^{(0)} = \nabla (r^L P_L(\cos \theta)) \left(\frac{e'}{(L+1)m} L + \mu' \right)$$

and when $L = 1$,

$$Q_{10}^{(0)} = \frac{e'}{2m} L + \mu'$$

in correspondence with the usual definition of the magnetic dipole moment.

Then (30.21) can be written

$$U_{i \rightarrow f} = (-i)^L \sqrt{\frac{\omega}{\pi}} \omega^L \sqrt{\frac{L+1}{L(2L+1)}} \frac{1}{(2L-1)!} Q_{LM}^{(0)}, \quad (30.23)$$

where $Q_{LM}^{(0)}$ denotes the matrix element

$$(Q_{LM}^{(0)})_{fi} = \int \varphi_f^* Q_{LM}^{(0)} \varphi_i d\tau.$$

We note that (30.23) differs from (30.13) only in that $Q_{LM}^{(1)}$ is replaced by $Q_{LM}^{(0)}$. The probability for emission of a photon with angular momentum L , projection of the angular momentum M , and parity $(-1)^{L+1}$ is thus given by

$$w_{LM}^{(0)} = \frac{2(L+1)}{L(2L+1)(2L-1)!} \omega^{2L+1} |Q_{LM}^{(0)}|^2, \quad (30.24)$$

which differs from (30.14) in that $Q_{LM}^{(1)}$ is replaced by $Q_{LM}^{(0)}$.

4. Selection Rules, Order of Magnitude Evaluations.

In order for photon emission to be possible, the initial and final electron states must satisfy certain definite selection rules following from the conservation laws for angular momentum and parity. They reduce simply to the fact that the parity in the final state must differ from that of the original one by a factor of $(-1)^L$ in the case of electric, and by a factor of $(-1)^{L+1}$ in the case of magnetic multipole radiation; the projections of the angular momenta of the initial m_1 and final m_2 states must be related by

$$m_1 - m_2 = M,$$

and the angular momenta j_1 and j_2 themselves must satisfy the inequality

$$|j_1 - j_2| \leq L \leq j_1 + j_2.$$

If these conditions are not satisfied the matrix element of the corresponding multipole moment vanishes.

For given j_1 and j_2 , the most probable emission is that of the photon with angular momentum

$$L = |j_1 - j_2|$$

(If this is compatible with the parity selection rule). Indeed the matrix elements (30.13) and (30.32) have integrands which depend on r in the combination $(\omega r)^L$. Since $\omega r \ll 1$,

$$Q_{L+2, M}^{(0)} \ll Q_{LM}^{(0)} \quad (\lambda = 0, 1).$$

(If $Q_{LM}^{(1)} \neq 0$, then $Q_{L+1, M}^{(0)} = 0$ according to the parity selection rule.)

Since the initial and final states are degenerate with respect to the projections of the angular momenta, we shall sum the probabilities over the initial, and shall average them over the final states:

$$w_L = \frac{1}{2L+1} \sum_{M, m_1, m_2} w_{LM} \delta_{m_1 - m_2, M}.$$

The nonzero values of the w_{LM} do not actually depend on M (due to spherical symmetry), and the summation therefore reduces to counting the number of transitions which are allowed by the selection rule. In particular, if $j_2 = 0$ (then $j_1 = L$), we obtain

$$w_L = w_{LM}.$$

The expressions for the transition probabilities as given by (30.14) and (30.23) contain the factor

$$\omega^{2L+1}.$$

If, therefore, the energy difference between the excited and ground state is small, and the angular momentum difference $j_1 - j_2$ is large, the transition probability can become very small, and the system may thus remain in the excited state for a relatively long time. Such states are called metastable.

In order to achieve a rough approximation of the absolute magnitude for the transition probability, we may replace the matrix element of the electric multipole moment, according to its definition (30.12) by eR^L :

$$Q_{LM}^{(0)} \sim eR^L, \quad (30.25)$$

where R gives the effective dimensions of the system. Similarly, for the magnetic multipole moment we can write, according to (30.22),

$$Q_{LM}^{(1)} \sim eR^L v, \quad (30.26)$$

where v is the ratio of the velocity of the charge to that light. We note that the order of magnitude of the velocity of electrons in atoms is given by

$$R\omega \sim v \sim v^2 = \frac{1}{137}.$$

For nucleons in nuclei, R , ω , and v have no definite relation.

Calculation of multipole moments according to (30.25) and (30.26) leads to the correct order of magnitude for the lifetimes of metastable nuclear states, which for values of L between 3 and 5 can attain values of the order of minutes, hours, and more. Nuclei in such states are called isomers, and behave almost like stable nuclei; they differ from nuclei which are in the ground state in several of their properties (for instance lifetimes for β -decay).

We have determined the probability for emission of photons in eigenstates of the angular momentum and parity. The results obtained can be used to find the probabilities for photon emission in momentum and polarization eigenstates. To do this we can expand the plane polarized wave which appears in the expression for the matrix element (30.2) in terms of spherical waves. Using (4.24), we have

$$e^{ikr} = \sum_{L, M; \lambda=0,1} \left(e_{\lambda} Y_{LM}^{(0)} \left(\frac{k}{\omega} \right) \right) a_{LM}^{(0)}. \quad (30.27)$$

If the wave length of the photon is large compared with the dimensions of the radiating system, one term of the sum in (30.27) is sufficient to determine the matrix element. Then the emission probability, denoted by $w_{k\lambda}$, can be written

$$w_{k\lambda} = |e_{\lambda} Y_{LM}^{(0)}|^2 w_{LM}^{(0)}, \quad (30.28)$$

where the $w_{LM}^{(0)}$ are given by (30.14) or (30.24). Equation (30.28) together with (4.25) gives the angular and polarization distribution of multipole radiation. Summing over the polarizations and integrating over the angle of emission, we obtain

$$\sum_{\lambda} \int w_{k\lambda} d\Omega = w_{LM}^{(0)}.$$

Thus far we have assumed that there are no photons in the initial state. If, however, there are N photons initially, the matrix element (30.1) contains a factor of $\sqrt{N+1}$ and the probability should therefore be multiplied by $N+1$.

We note in conclusion that although we have derived the equations for multipole radiation for a single particle in an external field, these equations are actually valid also for an arbitrary system of interacting particles so long as conditions $\omega R \ll 1$ and $\mathbf{v} \ll 1$ are satisfied. Then the multipole moment operators are the sums of these operators for the individual particles, and the wave functions ψ_1 and ψ_2 refer to the whole system. We shall not, however, bother proving this assertion.

5. Absorption of a Photon

Let us now consider photon absorption. The transition of an electron from state 1 to state 2, accompanied by the absorption of a photon from a state in which there are N photons, corresponds to the matrix element,

$$S_{1 \rightarrow 2}^{(1)} = -e \int \bar{\psi}_2 \hat{A} \psi_1 d^3x \sqrt{N}, \quad (30.29)$$

where the index 2 denotes the high-energy electron state, and 1 denotes the lower-energy state. Further, we shall use w_{21} to denote the probability of transition from state 2 to 1 with emission of a photon, and w_{12} to denote the probability of transition from 1 to 2 with absorption of a photon. It is easy to see, by comparing (30.29) and (30.1), that

$$w_{21} = \frac{N+1}{N} w_{12}. \quad (30.30)$$

If 2 is a state with angular momentum J_2 , and 1 is a state with angular momentum J_1 , then the probabilities averaged over the projections of the angular momenta are related by

$$w_{21} = \frac{N+1}{N} \frac{2J_1+1}{2J_2+1} w_{12}. \quad (30.31)$$

6. Photoelectric Effect.

If the photon energy is greater than the binding energy of the electron in the atom, the absorption of the photon causes ionization (the electron goes from a state belonging to the discrete spectrum to one in the continuous spectrum). This process is called the photoelectric effect.

Let us now obtain the general expression for the photoelectric-effect cross section. To do this we insert the potential of a photon in an eigenstate of the momentum and polarization into expression (30.29), obtaining

$$A = \frac{1}{\sqrt{2\omega}} e^{ikr}.$$

Integrating over the time, we obtain

$$S_{1 \rightarrow f}^{(1)} = -2\pi i U_{1 \rightarrow f} \sqrt{N} \delta(\epsilon_1 - \epsilon_2 + \omega),$$

$$U_{1 \rightarrow f} = -\frac{e}{\sqrt{2\omega}} \int \psi_2^* e \mathbf{A} \psi_1 e^{ikr} d^3r,$$

where ψ_1 and ψ_2 are time-independent wave functions. Using (28.18) to write an expression for the probability, and dividing it by the photon current density, which in this case is equal to unity, we obtain the following expression for the cross section for photon capture¹⁾:

$$\sigma = \frac{(2\pi)^3 e^2}{\omega} \left| \int \psi_2^* e \mathbf{A} \psi_1 e^{ikr} d^3r \right|^2 \delta(\epsilon_1 - \epsilon_2 + \omega). \quad (30.32)$$

In (30.32) the wave function of the final state is normalized so that

$$\int \psi_2^* \psi_2 d^3r = 1.$$

If we transform this to a wave function normalized for a unit energy interval, i.e., if we set

$$\int \psi_2^* \psi_2 d^3r = \delta(\epsilon - \epsilon'),$$

1) Here α is the fine structure constant, which in ordinary units is equal to $\frac{e^2}{4\pi\hbar c}$.

then

$$\sigma = \frac{(2\pi)^2 \hbar^2}{\omega} \left| \int \psi^* \mathbf{a} e^{i\mathbf{k} \cdot \mathbf{r}} \psi d\mathbf{r} \right|^2. \quad (30.33)$$

The problem reduces to calculating the interval in (30.33). We shall restrict ourselves to presenting, without the relevant calculations, several results referring to the photoelectric effect on the electrons of the K-shell of an atom¹⁾.

If the photon energy is large in comparison with the ionization potential, and if the nuclear charge is small, then we can neglect the effect of the nuclear Coulomb field on the final electron state; in this case ψ_2 may be considered a free-electron wave function (for instance, a plane wave). The total cross section for the photoelectric effect on the two electrons of the K-shell for arbitrary polarization and emission direction of the final electron state, is given by

$$\sigma = 4\pi r_0^2 Z^2 \alpha^4 \left(\frac{m}{\omega} \right)^8 (\eta^2 - 1)^2 \left[\frac{4}{3} + \frac{\eta(\eta-2)}{\eta+1} \left(1 - \frac{1}{2\eta \sqrt{\eta^2-1}} \ln \frac{\eta+1+\sqrt{\eta^2-1}}{\eta-1+\sqrt{\eta^2-1}} \right) \right], \quad (30.34)$$

where Z is the charge of the nucleus, $r_0^2 = \frac{\alpha}{m}$ and $\eta = \frac{\omega + m}{m}$.

In the ultrarelativistic case, when $\omega \gg m$, this equation can be written

$$\sigma = 4\pi r_0^2 Z^2 \alpha^4 \frac{m}{\omega} \left(\frac{m}{\omega} \right)^8, \quad \omega \gg m. \quad (30.35)$$

In the nonrelativistic limit, when $\omega \ll m$, (30.34) leads to

$$\sigma = \frac{32\sqrt{2}\pi}{3} r_0^2 \alpha^4 \left(\frac{m}{\omega} \right)^{1/2}. \quad (30.36)$$

We note that (30.36), as (30.34), is valid only when $\omega \gg \frac{1}{2} Z^2 \alpha^2 m$, where $\frac{1}{2} Z^2 \alpha^2 m$ is the ionization potential of the K-electron.

The cross section for the photoelectric effect with the nuclear Coulomb field taken into account can be obtained only in two limiting cases. In the nonrelativistic case, use of the exact nonrelativistic wave functions in (30.33) leads to (30.36) multiplied by

¹⁾ See W. Heitler, The Quantum Theory of Radiation (State Tech. Press, 1940).

$$f_{NR} = 2\pi \sqrt{\frac{T}{\omega-1}} \frac{e^{-4\kappa \arctan \xi}}{1 - e^{-2\kappa \xi}}, \quad (30.37)$$

where

$$\xi = \sqrt{\frac{T}{\omega-1}} = \frac{Z\alpha}{v}$$

(v is the velocity of the K-electron).

In the ultrarelativistic case, the result obtained differs from (30.36) by the factor,

$$f_{UR} = e^{-2\alpha} + 2(Z\alpha)^2 (1 - \ln 2\alpha). \quad (30.38)$$

In the following table we present the exact values for the cross section for the photoelectric effect, as obtained by numerical integration²⁾ for several cases in which the approximate formulas are not applicable.

Values of

$\frac{\sigma}{Z^2 \frac{8\pi}{3} r_0^2} (137)^4$	$\frac{\omega}{m}$		
Z	25	50	82
0.69	18	12.2	7.9
2.2	1.05	0.80	0.60

We also present the expression for the angular distribution of the photoelectrons, in the same approximation as is used for (30.36) ($\omega \ll m$, $\omega \gg \frac{1}{2}$)

$$\frac{d\sigma}{d\Omega} = \sigma \frac{3}{4\pi} \frac{\sin^2 \theta \cos^2 \varphi}{(1 - v \cos \theta)^4}, \quad (30.39)$$

where θ is the angle between the photon propagation vector $\mathbf{k}$ and the momentum $\mathbf{p}$ of the photoelectron, and φ is the angle between the plane defined by $\mathbf{p}$ and $\mathbf{k}$ and that defined by $\mathbf{k}$ and $\mathbf{e}$.

§ 31. Bremsstrahlung

1. General Expression for the Matrix Element.

In collisions involving an electron and another charge (or system of charges), in addition to electron scattering, photon emission may also take place. Such a process is called *bremsstrahlung* or *deceleration radiation*. If the electron collides with a heavy particle (a nucleus or an atom), the effect of the latter can be treated as an external field. Then the bremsstrahlung is described by the matrix element (30.1) in which the initial and final states belong to the continuous spectrum³⁾.

In view of the complex character of the electron wave functions in the field of the nucleus (see Sections 12, 13), the integral in (30.1) can be calculated analytically only for low energies⁴⁾ when the nonrelativistic approximation may be used.

²⁾ H. R. Hulme, J. Mc. Donnell, R. Buckingham and R. Fowler, Proc. Roy. Soc. 149, 131 (1935).

³⁾ The final state may belong to the discrete spectrum. We shall not consider this case.

⁴⁾ Recently results for the high energy limit ($\epsilon \gg m$) and small electron-scattering angles have also been obtained. See Maximov, and Bethe, Phys. Rev. 87, 158 (1952); Davies and Bethe, loc. cit.; Bethe, Maximov, and Low, Phys. Rev. 91, 417 (1953).

We shall present these results somewhat later, first considering another somewhat simpler case which has a wide range of applications.

Let us assume that the external field is such that it can be treated by perturbation methods (the criterion for the applicability of perturbation theory is clearly the same as that for the Born approximation for a Coulomb field, namely $Z e^2 \gamma_{\text{ext}} \ll 1$). Then to determine the emission probability it is sufficient to calculate the appropriate element of the second-order matrix $S^{(2)}$ of (29.2). This matrix element will differ from (29.3) for scattering of a photon by an electron only in that the potential of the incident photon is replaced by that of the external field, so that

$$S^{(2)}_{if} = \frac{e^2}{2} \int \bar{\psi}_2(2) [\hat{A}^*(2) S^P(21) \hat{A}^{(0)}(1) + \hat{A}^{(0)}(2) S^P(21) \hat{A}^*(1)] \psi_1(1) \cdot d^4x_2 d^4x_1. \quad (31.1)$$

Here ψ_1 and ψ_2 are the initial and final free electron wave functions, $\hat{A}$ is the potential of the emitted photon, and $\hat{A}^{(0)}$ is the external field potential.

2. Perturbation Theory for the Wave Function of an Electron in the Continuous Spectrum.

By comparing the second-order matrix element (31.1) with the first-order matrix element (30.1), we can find the wave function of the electron in the external field (in first-order perturbation theory). Indeed, (31.1) should be identical with (30.1) if the wave functions in the latter equation take account of scattering in the external field. Let us denote these, as opposed to the free-electron wave functions (ψ_1 and ψ_2) by ψ_1^e and ψ_2^e and let us write

$$\psi_s^e = \psi_s + \psi_s^e \quad (s=1, 2).$$

Inserting this into (30.1) and noting that for free-electron wave functions (30.1) vanishes, comparison with (31.1) leads to

$$\psi_2^{(0)}(1) = \frac{e}{2} \int \bar{\psi}_2(12) \hat{A}^{(0)}(2) \psi_1(2) d^4x_2. \quad (31.2)$$

Similarly,

$$\bar{\psi}_1^{(0)}(1) = \frac{e}{2} \int \bar{\psi}_1(2) \hat{A}^{(0)}(2) S^P(21) d^4x_2.$$

If $\psi = u e^{i p x}$, then

$$\psi^e = u e^{i p x} - i e \int \frac{d^4 p'}{f^2 + m^2} \hat{A}_{f-p'} u e^{i f' x} d^4 f'$$

[where $\hat{A}_1$ represents the Fourier component of the external field, to be defined below; see (31.6)]. In an electrostatic field, using (31.7), we have

$$\psi^e = u e^{i p x} + e 2\pi \int \frac{d^4 p'}{f^2 + m^2} \hat{A}_{f-p} u e^{i f' x} d^4 f'.$$

In the nonrelativistic limit even small values of $|f|$ are significant, so that $\hat{f} = -\frac{m}{2m}$ and

$$\psi^e = \psi_p + e \int \frac{d^4 f - p}{e_p - e_f} \psi_f d^4 f \quad (31.3)$$

$$(e_p - e_f = \frac{p^2}{2m} - \frac{f^2}{2m});$$

This is identical with the expression for the wave function of the continuous spectrum in nonrelativistic perturbation theory¹⁾. Here, as follows from the general definition of $\hat{S}^E$, the integration in (31.3) should be performed according to the rules previously described for the poles of the integrand [see (18.34)].

3. Bremsstrahlung Cross Section

Inserting the electron wave functions and photon potentials given by (29.4) and Equation (18.33) for $\hat{S}^E$ into (31.1), we obtain the following equation for the matrix element [which can also be written on the basis of the general rules for writing out the S matrix in momentum space; see (24.7)]:

$$i S^{(2)}_{if} = -\frac{i e^2}{2\omega} u_2 \left\{ \hat{e} \frac{i \hat{f}_1 - m}{m^2 \epsilon_1} \hat{a}_q + \hat{a}_q \frac{i \hat{f}_2 - m}{m^2 \epsilon_2} \right\} u_1, \quad (31.4)$$

where

$$\left. \begin{aligned} f_1 &= p_2 + k, & f_2 &= p_1 - k, \\ q &= p_2 + k - p_1, \\ m^2 \epsilon_1 &= f_1^2 + m^2 = 2 p_2 k = -2 \epsilon_2 \omega (1 - \cos \theta_2), \\ m^2 \epsilon_2 &= f_2^2 + m^2 = -2 p_1 k = 2 \epsilon_1 \omega (1 - \cos \theta_1), \end{aligned} \right\} \quad (31.5)$$

(θ_1, θ_2 are the angles between $\mathbf{k}$ and $\mathbf{p}_1, \mathbf{p}_2$), and

$$a_q = a(q) = \int A^{(0)} e^{-i q x} d^4 x. \quad (31.6)$$

Figure 35 shows the diagrams corresponding to the matrix element (31.4). The designation and order of the lines on these diagrams are the same as in Fig. 32.

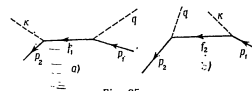


Fig. 35

The matrix element (31.4) and Fig. 35 which illustrates it, has the same structure as (29.5) (Fig. 32). Corresponding to the scattered photon with momentum $\mathbf{k}$ and polarization $\hat{e}_2$ in (29.5) is the photon here emitted with momentum $\mathbf{k}$ and polarization $\hat{e}$. Corresponding to the first ($\mathbf{p}_1, \mathbf{e}_1$) photon is the "pseudophoton" with four-momentum $\mathbf{g}$ and amplitude $\hat{a}_g$. The meaning of the vector $\mathbf{g}$ is the following: the negative of its space part $-\mathbf{g}$ is the momentum, and $\frac{1}{2}\mathbf{g}$ is the energy, transferred to the nucleus (or atom).

¹⁾ See L. Landau and E. Lifshits, Quantum Mechanics (State Tech. Press, 1948) Equation (49.3).

If the external field is time independent and given by a static potential (infinitely heavy nucleus), then

$$q_4 = 0, \quad a_q = i a_q^0 \cdot 2\pi\delta(\epsilon_2 + \omega - \epsilon_1), \quad (31.7)$$

where

$$i a_q^0 = \int A^{(0)} e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (31.8)$$

(the polarization of the pseudophoton, as opposed to that of the real one, is directed along the $\mathbf{x}_4$ axis).

For the Coulomb field of a nucleus with charge $-Ze$

$$A^{(0)} = -\frac{Ze}{q^2}. \quad (31.9)$$

Let us now find the differential cross section for emission of a photon by an electron in an electrostatic field. This can be done from the formula

$$d\sigma = \frac{e^4}{2\omega} |a_q^0|^2 |\bar{u}_2 Q u_1|^2 dP_2 d\mathbf{k} \frac{\delta(\epsilon_1 - \epsilon_2 - \omega)}{(2\pi)^3 J},$$

where according to (31.4) and (31.7)

$$Q = \frac{1}{m^2 \epsilon_1} \hat{\epsilon} (\hat{p}_1 - m) \beta + \frac{\beta}{m^2 \epsilon_2} (\hat{p}_2 - m) \hat{\epsilon}.$$

In this case the current density $\mathbf{J}$ coincides with the initial electron velocity $\mathbf{v}_1$. Since here, as opposed to the case of photon scattering, $\mathbf{p}_2$ and $\mathbf{k}$ are independent, we have

$$d\mathbf{k} dP_2 \delta(\epsilon_1 - \epsilon_2 - \omega) \rightarrow d\mathbf{k} \frac{P_2^2}{v_2} d\alpha_2$$

(where $d\alpha_2$ is the element of solid angle about $\mathbf{p}_2$).

Averaging over the initial electron spins and summing over the final electron spins and photon polarizations, we obtain

$$d\sigma = e^4 |a_q^0|^2 \frac{1}{4\omega} \frac{1}{P_1} \frac{1}{4} \text{Sp} \frac{d\mathbf{k} d\alpha_2}{(2\pi)^3}, \quad (31.10)$$

where, in analogy with (29.18) and (29.19),

$$\begin{aligned} &= \left[\frac{1}{m^2 \epsilon_1} \gamma_1 (\hat{p}_1 - m) \beta + \frac{1}{m^2 \epsilon_2} \beta (\hat{p}_2 - m) \gamma_1 \right] (\hat{p}_1 - m) \times \\ &\quad \times \left[\frac{1}{m^2 \epsilon_1} \beta (\hat{p}_1 - m) \gamma_1 + \frac{1}{m^2 \epsilon_2} \gamma_1 (\hat{p}_2 - m) \beta \right] (\hat{p}_2 - m). \end{aligned} \quad (31.11)$$

The calculation of Sp F may be performed as in Section 29. First, Equations (7.27) and (28.12) are used to eliminate γ_1 matrices, and then the commutation rules are used to bring two factors of β together. This latter is very simple to achieve, since every four-vector $\mathbf{b}$ whose space components are real and whose fourth component is imaginary satisfies

$$\beta \hat{\mathbf{b}} = -\hat{\mathbf{b}}^* \beta.$$

After eliminating β (noting that $\beta^2 = 1$) each term contains products of no more than four matrices.

4. Angular Distribution for Radiation in a Coulomb Field.

We present here the final result¹⁾ for the bremsstrahlung differential cross section in a Coulomb field, when $\frac{q^2}{m^2}$ is given by (31.9):

$$d\sigma = \frac{Z^2 e^3}{(2\pi)^2} \frac{d\omega}{\omega} \frac{1}{P_1} \frac{d\alpha_2}{q^4} \frac{4}{m^4} \left\{ 4 \left[\mathbf{k}_1, \frac{\epsilon_1 \mathbf{p}_2}{\epsilon_1} + \frac{\epsilon_2}{\epsilon_2} \mathbf{p}_1 \right]^2 - \right. \\ \left. - q^2 \left[\mathbf{k}, \frac{\mathbf{p}_2 + \mathbf{p}_1}{\epsilon_1 \epsilon_2} \right]^2 - \frac{2\omega^2}{\epsilon_1 \epsilon_2} [\mathbf{k}, \mathbf{p}_2 - \mathbf{p}_1]^2 \right\}. \quad (31.12)$$

Here $\mathbf{q}$, ϵ_1 and ϵ_2 are functions of $\mathbf{k}$, $\mathbf{p}_1$ and $\mathbf{p}_2$ given by (31.5), and $\alpha = \frac{q^2}{4\epsilon^2} = \frac{1}{137}$.

In general the angular distribution of the radiation is quite complex. We shall first consider the case of low initial electron energies $|\mathbf{p}_1| \ll m$, which corresponds to the problem of the continuous x-ray spectrum. Since in this case

$$\omega = \frac{P_1^2 - P_2^2}{2m} \ll |\mathbf{p}_1|,$$

we have

$$q^2 = (P_1 - P_2)^2$$

which is independent of $\mathbf{k}$. Further,

$$m^2 \epsilon_2 = -m^2 \epsilon_1 = 2m\omega, \quad \epsilon_1 = \epsilon_2 = m.$$

We need keep only the first term in the braces of (31.12). We then get

¹⁾ H. Bethe and W. Heitler, Proc. Roy. Soc. 146, 83 (1934).

$$d\sigma = \frac{Z^2 e^4}{\pi^2} \frac{d\omega}{\omega} \frac{|p_2|}{|p_1|} \frac{d\theta_2}{(p_1 - p_2)^4} \frac{(k \cdot p_1 - p_2)^2}{\omega^2} \quad (31.13)$$

Equation (31.13) has a very simple structure. It can be approximated (noting that $|p_2|$ is not very different from $|p_1|$) by the product of the Rutherford cross section, for elastic electron scattering

$$\frac{(2me^2)^2}{(p_1 - p_2)^4} d\omega_2$$

and the probability for radiation during scattering

$$\frac{\alpha}{(2\pi)^3} \left[\frac{k}{\omega} \cdot v_1 - v_2 \right]^2 \frac{d\omega}{\omega} d\omega_2.$$

This last equation is in agreement with the classical formula for low-frequency dipole radiation¹⁾ (see Section 32). The maximum intensity of the radiation is emitted in a direction perpendicular to the plane of motion of the electron, which is defined by p_1 and p_2 .

Let us now consider the opposite limit of high electron energies, $\epsilon_1 \gg m$. Since the electron velocity is close to that of light, the occurrence of the factors

$$x_1 \sim (1 - v_2 \cos \theta_2); \quad x_2 \sim (1 - v_1 \cos \theta_1)$$

in the denominator of (31.12) causes the differential cross section to have a sharp maximum for $\theta_1 = 0$, $\theta_2 = 0$, that is, in the initial direction of motion of the electron. The radiation lies within a solid angle of order m/ϵ about this direction

5. Bremsstrahlung Spectrum

Let us now find the total cross section for radiation of a photon with a given frequency independent of the directions of k and p_2 . Let us denote this cross section by $d\sigma_\omega$. Integration of (31.12) over the angles gives

$$d\sigma_\omega = \frac{\alpha}{\omega} \frac{d\omega}{\omega} \frac{|p_2|}{|p_1|} \left[\frac{4}{3} - 2\epsilon_1 \epsilon_2 \frac{p_1^2 + p_2^2}{p_1^2 p_2^2} + \right. \\ \left. + m^2 \left(\frac{\lambda_1 \epsilon_2}{|p_1|^3} + \frac{\lambda_2 \epsilon_1}{|p_2|^3} - \frac{\lambda_1 \lambda_2}{|p_1| |p_2|} \right) + \left(\frac{8}{3} \frac{\epsilon_1 \epsilon_2}{|p_1| |p_2|} + \frac{\omega^2}{|p_1|^3 |p_2|^3} (\epsilon_1^2 \epsilon_2^2 + p_1^2 p_2^2) + \right. \right. \\ \left. \left. + \frac{m^2 \omega}{2 |p_1| |p_2|} \left(\frac{\epsilon_1 \epsilon_2 + p_1^2}{|p_1|^3} \lambda_1 - \frac{\epsilon_1 \epsilon_2 + p_2^2}{|p_2|^3} \lambda_2 + \frac{2\omega \epsilon_1 \epsilon_2}{p_1^2 p_2^2} \right) \right) \right] \quad (31.14)$$

¹⁾ See L. Landau and E. Lifshits, Field Theory (State Tech. Press, 1948).

where

$$L = \ln \frac{p_1^2 + |p_1| |p_2| - \epsilon_1 \omega}{p_1^2 - |p_1| |p_2| - \epsilon_1 \omega} = 2 \ln \frac{\epsilon_1 \epsilon_2 + |p_1| |p_2|}{\epsilon_1 \omega} - m^2,$$

$$\lambda_1 = \ln \frac{\epsilon_1 + |p_1|}{\epsilon_1 - |p_1|} = 2 \ln \frac{\epsilon_1 + |p_1|}{m},$$

$$\lambda_2 = 2 \ln \frac{\epsilon_2 + |p_2|}{m},$$

$$\bar{\Phi} = Z^2 r_0^2 \alpha.$$

For small energies ($|p_1| \ll m$) (31.14) takes on the simple form [this expression can also be obtained by integrating (31.13)]

$$d\sigma_\omega = \frac{16}{3} \bar{\Phi} \frac{d\omega}{\omega} \frac{m^2}{p_1^2} \ln \frac{|p_1| + |p_2|}{|p_1| - |p_2|} = \frac{8}{3} \bar{\Phi} \frac{d\omega}{\omega} \frac{m}{T_1} \ln \frac{(\sqrt{T_1} + \sqrt{T_1 - \omega})^2}{\omega}, \quad (31.15)$$

$$|p_1| \ll m,$$

where

$$T_1 = \frac{p_1^2}{2m} = \epsilon_1 - m$$

is the initial kinetic energy of the electron.

It can be seen from (31.15) that the cross section for radiation is approximately inversely proportional to the frequency. For the maximum possible frequency $\omega = T_1$,

$$d\sigma_\omega = 0.$$

As $\omega \rightarrow 0$, the intensity $\omega d\sigma_\omega$ diverges logarithmically. This divergence is strongly related to the divergence of the Rutherford formula for small angles, and characterizes the pure Coulomb field. When the screening of the nuclear field by the atomic electrons is taken into account, this divergence is liquidated as in the case of elastic scattering. We shall return to the question of screening somewhat later.

For low energies, as was mentioned at the beginning of this section, (30.1) can be calculated using the exact nonrelativistic electron wave functions²⁾.

The result of this calculation reduces to multiplying (31.13) and (31.15) by f , where

$$f = \frac{(2\pi)^2 \epsilon_1 \epsilon_2}{(1 - e^{-2\pi k}) (\epsilon_1^2 \epsilon_2^2 - 1)}, \quad (31.16)$$

²⁾ A. Sommerfeld, Ann. Physik 11, 257 (1931).

In which

$$\xi_{1,2} = \frac{Za}{v_{1,2}}$$

When $\xi_1 \ll 1$ and $\xi_2 \ll 1$, (31.16) becomes unity, as is to be expected since then the Born approximation becomes applicable. In the short wavelength limit of the spectrum (when $p_2 \rightarrow 0$), however $\xi_2 \rightarrow \infty$, and $\bar{f}$ becomes infinite, independent of the initial velocity. Therefore at the boundary of the spectrum $\omega = T_1$, the cross section $d\sigma$ remains finite.

In the high energy limit, when $\epsilon_1 \gg m$ and $\epsilon \gg m$ (31.14) becomes

$$d\sigma_\omega = 4\bar{\Phi} \frac{d\omega}{\omega} \frac{\epsilon_2}{\epsilon_1} \left[\frac{\epsilon_1^2 + \epsilon_2^2}{\epsilon_1 \epsilon_2} - \frac{2}{3} \right] \left[\ln \frac{2\epsilon_1 \epsilon_2}{m^2} - \frac{1}{2} \right] = \quad (31.17)$$

$$= 4\bar{\Phi} \frac{d\omega}{\omega} \left[1 + (1-u)^2 - \frac{2}{3}(1-u) \right] \left[\ln \left(\frac{2\epsilon_1}{m} \frac{1-u}{u} \right) - \frac{1}{2} \right],$$

where $u = \frac{\omega}{\epsilon_1}$.

We note that the probability that the electron radiates a given fraction of its energy (that is, the radiation probability for a given value of ω/ϵ_1) increases approximately as the logarithm of ϵ_1/m . For low frequencies (31.17) is essentially inversely proportional to ω , and $\omega d\sigma_\omega$ diverges logarithmically as $\omega \rightarrow 0$, just as in the nonrelativistic case.

Let us also give the expression for the cross section for electron energy loss due to radiation, that is for the quantity

$$\Phi = \frac{1}{v_1} \int_0^{\epsilon_1 - m} \omega d\sigma_\omega.$$

Inserting $d\sigma_\omega$ from (31.14), we obtain

$$\Phi = \bar{\Phi} \left\{ \frac{12\epsilon_1^2 + 4m^2}{3\epsilon_1 |p_1|} \ln \frac{\epsilon_1 + |p_1|}{m} - \frac{(8\epsilon_1 + 6|p_1|)m^2}{3\epsilon_1 p_1^2} \left(\ln \frac{\epsilon_1 + |p_1|}{m} \right)^2 - \frac{4}{3} + \right. \quad (31.18)$$

$$\left. + \frac{2m^2}{\epsilon_1 |p_1|} F \left(\frac{2|p_1|(\epsilon_1 + |p_1|)}{m^2} \right) \right\},$$

where F is the function defined by

$$F(x) = \int_0^x \frac{\ln(1+y)}{y} dy.$$

For $x \ll 1$,

$$F(x) = x - \frac{x^2}{4} + \frac{x^3}{9} - \frac{x^4}{16} + \dots$$

For $x \gg 1$ we can use the same series for $F(1/x)$, and the relation

$$F(x) + F\left(\frac{1}{x}\right) = \frac{\pi^2}{6} + \frac{1}{2}(\ln x)^2.$$

For low energies, when $|p_1| \ll m$, it follows from (31.18) that

$$\Phi = \frac{16}{3} \bar{\Phi}, \quad (31.19)$$

and for very high energies, when $\epsilon_1 \gg m$,

$$\Phi = 4 \left(\ln \frac{2\epsilon_1}{m} - \frac{1}{3} \right) \bar{\Phi}. \quad (31.20)$$

Thus for low energies, the ratio of the mean energy radiated to the initial energy is constant. For high energies, this ratio increases in proportion to the logarithm of the energy.

6. Screening

All the above results from (31.12) on, are obtained for the Coulomb field, whose Fourier components are given by (31.9). Since nuclei are ordinarily surrounded by electrons, the nuclear field is Coulombic only for distances smaller than the radius of the K-shell; at large distances the field is partially or entirely screened. In order to clarify the role of distance from the nucleus in radiation processes, let us consider the general expression for the Fourier component A_q^i as given by (31.8):

$$iA_q^i = \int A^{(0)} e^{-iqr} dr.$$

The region which essentially determines the value of this integral is given by the inequality

$$r \lesssim \frac{1}{|q|},$$

since for larger values of r the oscillations of e^{-iqr} become important. The minimum value of q which corresponds to the maximum effective distance, namely $q_{\min}$, can be found from the conservation laws. For low energies and small distances ($|p_1| \ll m$; $\omega \ll |p_1|$)

$$q_{\min} \approx \frac{m}{|p_1|} \omega.$$

For large energies (ϵ_1 and $\epsilon_2 \gg m$)

$$q_{\min} \approx \frac{m^2 \omega}{2\epsilon_1 \epsilon_2} \quad (31.21)$$

It is seen from this that for sufficiently low frequencies $\frac{1}{q_{\min}}$ can be greater than the dimensions of the atom, and the probability for emission of such photons will be much less than that calculated according to (31.14). Therefore when $\omega \rightarrow 0$, the product $\omega d\sigma_{\omega}$ vanishes, which differs from its behavior in a pure Coulomb field.

For sufficiently high energies, $q_{\min}$ becomes small even for frequencies of the order of the original electron energy. Indeed if ω and ϵ_2 are of order $\epsilon_1 \gg m$, then according to (31.21)

$$\frac{1}{q_{\min}} \approx \frac{\epsilon_1}{m^2}$$

When $\epsilon_1 \approx 137 \frac{m}{Z}$, the quantity $\frac{1}{q_{\min}}$ becomes equal to the radius of the K-shell, and therefore screening must be taken into account. When $\epsilon_1 > 137 \frac{m}{Z}$, the quantity $\frac{1}{q_{\min}}$ becomes greater than the effective radius of the atom, which, according to the Thomas-Fermi model is $\frac{a_0}{Z^{1/3}}$, where a_0 is the radius of the hydrogen atom.

We present here the expression for $d\sigma_{\omega}$ with screening taken into account ($q_{\min}^1$ is calculated according to the Thomas-Fermi atomic potential¹⁾):

$$d\sigma_{\omega} = Z^2 a_0^2 \frac{1}{\epsilon_1^2} \frac{d\omega}{\omega} \left[(\epsilon_1^2 + \epsilon_2^2) \left(\Phi_1(\zeta) - \frac{4}{3} \ln Z \right) - \frac{2}{3} \epsilon_1 \epsilon_2 \left(\Phi_2(\zeta) - \frac{4}{3} \ln Z \right) \right], \quad (31.22)$$

where

$$\zeta = 100 \frac{m\omega}{\epsilon_1 \epsilon_2 Z^{1/3}},$$

where Φ_1 and Φ_2 are the functions shown in Fig. 36 in the interval $0 < \zeta < 20$. When $\zeta \ll 1$ (low frequency or high energy) we have so-called complete screening. In this case

$$\Phi_1(0) = 4 \ln 183, \quad \Phi_2(0) = \Phi_1(0) - \frac{2}{3}$$

and

$$d\sigma_{\omega} = Z^2 a_0^2 \frac{1}{\epsilon_1^2} \frac{d\omega}{\omega} 4 \left[\left(\epsilon_1^2 + \epsilon_2^2 - \frac{2}{3} \epsilon_1 \epsilon_2 \right) \ln 183 Z^{-1/3} + \frac{\epsilon_1 \epsilon_2}{9} \right]. \quad (31.23)$$

¹⁾ See the reference on p. 319.

When $\zeta > 1$, we return to the case of no screening as given by (31.17).

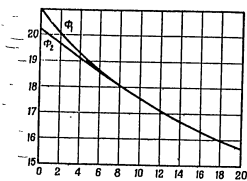


Fig. 36

When $2 < \zeta < 15$ Equation (31.17) can still be used if $\frac{1}{2}$ is replaced by $\frac{1}{2} + \zeta$ in the last factor, where ζ is given by the table below:

ζ	2	2.5	3	4	5	6	8	10	15
$\zeta(\zeta)$	0.21	0.16	0.13	0.09	0.065	0.05	0.03	0.02	0.01

The cross section for energy losses to radiation in the case of complete screening is

$$\Phi = \bar{\Phi} \left[4 \ln (183 Z^{-1/3}) + \frac{2}{9} \right]. \quad (31.24)$$

We see that the screening eliminates the logarithmic increase of the cross section with energy which was given by (31.20).

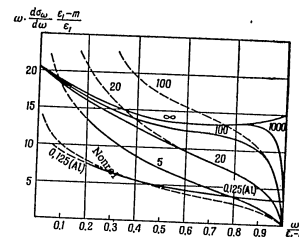
Figure 37 shows the dependence of $\omega \frac{d\sigma_{\omega}}{d\omega} \frac{\epsilon_1 - m}{\epsilon_1 m}$ on $\frac{\omega}{\epsilon_1 m}$ for various values of $\frac{\epsilon_1 - m}{\epsilon_1 m}$ (the latter are given by the numbers on the curves). These curves include screening as well as the deviation from the Born approximation (for low energies) in the region of the low-wavelength limit. Figure 38 shows the dependence of $\frac{\Phi}{\bar{\Phi}}$ on $\frac{\epsilon_1 - m}{m}$ for various substances with screening included. The highest curve represents absence of screening.

7. Radiation from Electron-Electron and Electron-Positron Collisions.

Let us now consider photon emission arising from electron-electron collisions. In this problem we may no longer replace the effect of one of the particles by an external field. It is therefore necessary to consider those S matrix elements involving one photon and four electron (two initial and two final) states. These are contained in the third-order matrix $S^{(3)}$. We shall calculate the matrix element directly from the general rules formulated in Sections 23 and 24.

Figure 39 shows diagrams corresponding to individual terms of this matrix element. Altogether there are eight diagrams. Next to each line segment we have indicated the appropriate four-momenta; p_1 and p_1' are the initial electron momenta, p_2 and p_2' are the final electron momenta, and k is the momentum of the emitted

¹⁾ See the reference on p. 369.



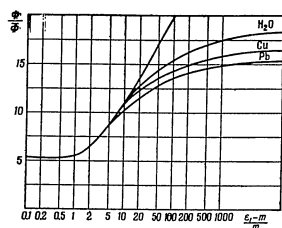


Fig. 38

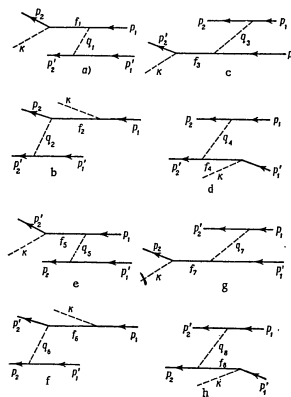


Fig. 39

photon; the f_i are the four-momenta of the virtual electrons, and the q_i are those of the virtual photons ($i = 1, 2, \dots, 8$). The last four diagrams (e, f, g, h) differ from the first four (a, b, c, d) by the interchange of p_2 and p_1' ("exchange" diagrams). The terms corresponding to these enter the matrix elements with opposite signs. In diagrams a and c the scattering takes place "first" and the radiation "second," and in diagrams b and d, vice versa. In diagrams a and b "the first electron radiates" and in diagrams c and d, the second one does. The values of f_i and q_i are easily obtained from the conservation laws at each of the vertices. Thus,

$$\begin{aligned} q_1 &= q_2 = p_2' - p_1', & q_3 &= q_4 = p_2 - p_1, \\ f_1 &= p_2 + k, & f_2 &= p_1 - k, \\ f_3 &= p_2' + k, & f_4 &= p_1' - k. \end{aligned} \quad (31.25)$$

Moving along the electron lines, and replacing the line segments and vertices by the appropriate amplitudes and operators, we obtain the following expression for the matrix element (in each term the first factor refers to the upper electron line, and the second one, to the lower one):

$$\begin{aligned} S_{fi}^{(0)} &= \frac{e^4}{\sqrt{2\omega}} \left\{ \left(\bar{u}_2 \hat{\epsilon} \frac{f_1 - m}{f_1^2 + m^2} \gamma_4 u_1 \right) \frac{1}{q_1^2} (\bar{u}_2' \gamma_1 u_1') + \right. \\ &+ \left(\bar{u}_2' \gamma_1 \frac{f_2 - m}{f_2^2 + m^2} \hat{\epsilon} u_1 \right) \frac{1}{q_2^2} (\bar{u}_2' \gamma_1 u_1') + (\bar{u}_2' \gamma_1 u_1') \frac{1}{q_3^2} \left(\bar{u}_2 \hat{\epsilon} \frac{f_3 - m}{f_3^2 + m^2} \gamma_4 u_1' \right) + \\ &+ (\bar{u}_2' \gamma_1 u_1') \frac{1}{q_4^2} \left(\bar{u}_2 \gamma_1 \frac{f_4 - m}{f_4^2 + m^2} \hat{\epsilon} u_1 \right) - \left(\bar{u}_2 \hat{\epsilon} \frac{f_5 - m}{f_5^2 + m^2} \gamma_4 u_1 \right) \frac{1}{q_5^2} (\bar{u}_2' \gamma_1 u_1') - \\ &- \left(\bar{u}_2' \gamma_1 \frac{f_6 - m}{f_6^2 + m^2} \hat{\epsilon} u_1 \right) \frac{1}{q_6^2} (\bar{u}_2' \gamma_1 u_1') - (\bar{u}_2' \gamma_1 u_1') \frac{1}{q_7^2} \left(\bar{u}_2 \hat{\epsilon} \frac{f_7 - m}{f_7^2 + m^2} \gamma_4 u_1' \right) - \\ &- (\bar{u}_2' \gamma_1 u_1') \frac{1}{q_8^2} \left(\bar{u}_2 \gamma_1 \frac{f_8 - m}{f_8^2 + m^2} \hat{\epsilon} u_1' \right) \left. \right\} (2\pi)^4 \delta(p_1 + p_1' - p_2 - p_2' - k). \end{aligned} \quad (31.26)$$

The expression for the cross section corresponding to the matrix element (31.26) is extremely complicated (after summing over electron spins and photon polarizations, there arise many terms containing the trace of a product of six γ -matrices). We shall restrict ourselves to presenting the results for two limiting cases: the nonrelativistic

$$|p_1|, |p_1'| \ll m$$

and the ultrarelativistic

$$|p_1| \gg m.$$

In the nonrelativistic case the cross section can also be obtained by using Equation (30.14) for electric quadrupole radiation (the dipole moment of a two-electron system is zero)¹⁾. This can be done by using the fact that in the nonrelativistic approximation a system of two particles can be replaced by a single particle in an

1) In Section 30 we made use of the approximation in which $\omega \ll 1$, where ω represents significant distances for a given problem. For bound states, ω represents atomic (nuclear) dimensions. In the present case, $\omega \sim \gamma v$, where v is the velocity of the electron, and τ is the "collision time" given by $\tau \sim \frac{1}{\omega}$. Thus, the approximation of multiple radiation and the nonrelativistic limit agree in that they require $\gamma \ll 1$.

external field (the Coulomb interaction field for electrons). Then the wave functions of the initial and final states can be chosen according to (31.3), in which

$$\psi_p = \frac{1}{\sqrt{2}} (e^{i\mathbf{p}\cdot\mathbf{r}} \pm e^{-i\mathbf{p}\cdot\mathbf{r}}) e^{-i\epsilon t},$$

where $\mathbf{r}$ is the relative position vector, and $\mathbf{p}$ is the relative momentum (the + sign corresponds to total spin 0, and the - sign corresponds to total spin 1).

The expression for the differential cross section can be written¹⁾

$$d\sigma = \frac{16}{45} \alpha r_0^2 \left\{ \frac{3}{q^2} [4(p_1^2 - p_2^2)^2 + 3(p_1 p_2)^2] + \frac{1}{q^4} [144(p_1 p_2)^4 + \right. \\ \left. + 264(p_1^2 - p_2^2)^2(p_1 p_2)^2 + 105(p_1^2 - p_2^2)^4(p_1 p_2)^2 + 12(p_1^2 - p_2^2)^6] - \right. \\ \left. - \frac{1}{q^6} [36(p_1 p_2)^4 + 39(p_1^2 - p_2^2)^2(p_1 p_2)^2 + 12(p_1^2 - p_2^2)^4] \right\} \frac{d\mathbf{p}}{|p_1(p_1^2 - p_2^2)|}, \quad (31.27)$$

$$\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2; \quad s = \mathbf{p}_1 + \mathbf{p}_2.$$

The energy of the emitted photon is

$$\omega = \epsilon_1 - \epsilon_2 = \frac{p_1^2}{m} - \frac{p_2^2}{m},$$

where ϵ_1 and ϵ_2 are the initial and final electron energies in the center-of-mass system.

The angular distribution of the radiation in the center-of-mass system of two electrons is determined by its quadrupole character ($\sigma \propto |\mathbf{E}_\perp(\theta)|^2$).

Integration of (31.27) over the angles, which should be performed only over a hemisphere in view of the indistinguishability of the particles, gives

$$d\sigma_\omega = \frac{4}{15} \alpha r_0^2 \left\{ 17 - \frac{3x^2}{(2-x)^2} + \frac{12(2-x)^4 - 7(2-x)^2 x^2 - 3x^4}{(1-x)^2(2-x)^2} \times \right. \\ \left. \times \ln \left[\frac{1}{\sqrt{x}} + \sqrt{\frac{1}{x} - 1} \right] \right\} \frac{\sqrt{1-x} dx}{x}, \quad (31.28)$$

where

$$x = \frac{\omega}{\epsilon_1}.$$

1) E. Lifshits, J. Exptl.-Theoret. Phys. 19, 562 (1948); B. Fedyushin, J. Exptl.-Theoret. Phys. 22, 140 (1952).

The cross section for energy loss due to radiation by an electron colliding with an electron at rest is given by the equation

$$\Phi = \frac{1}{\epsilon_1} \int \omega d\sigma_\omega \approx 8\alpha r_0^2, \quad (31.29)$$

which of the same order as that obtained in the field of the nucleus for $Z = 1$ (see 31.19).

Since the atom of the nucleus whose charge is $-Ze$ has Z electrons, the energy loss due to radiation by an electron colliding with electrons in the shells is $1/Z$ of the losses due to collisions with the nucleus ($\Phi \approx Z^{-1}$).

In the ultrarelativistic case the cross section for radiation due to collision with a stationary electron is given by¹⁾

$$d\sigma_\omega = 2\alpha r_0^2 \frac{\epsilon_1}{\epsilon_2} \frac{d\omega}{\omega} \left\{ \left(-\frac{2}{3} + \frac{\epsilon_1^2 + \epsilon_2^2}{\epsilon_1 \epsilon_2} \right) \left(2 \ln \frac{2\epsilon_1 \epsilon_2}{\omega m} - 1 \right) - \left(1.70 \frac{\omega^2}{|p_1| |p_2|} + 2.14 \right) \right\}, \quad (31.30)$$

$$\Phi = 4\alpha r_0^2 \left(\ln \frac{2\epsilon_1}{m} - 1.05 \right).$$

Up to a factor in the logarithm, the cross section $d\sigma_\omega$ is the same as that for radiation by an electron in the field of a nucleus with $Z = 1$. This result is easy to understand on the basis of the following considerations. As we have seen [see (31.21)] for high energies the efficiency of momentum transfer to the nucleus becomes extremely small, and therefore the mass of the nucleus becomes less and less important.

For collision of an electron with a positron, the ultrarelativistic case gives exactly the same result as for electron-electron collision. In the nonrelativistic case, however, dipole radiation can take place. In this case the cross section for radiation is

$$d\sigma_\omega = \alpha r_0^2 \frac{32}{3} \frac{m d\omega}{\epsilon_1 \omega} \ln \frac{1 + \sqrt{1 - \frac{2m}{\epsilon_1}}}{1 - \sqrt{1 - \frac{2m}{\epsilon_1}}}. \quad (31.31)$$

This last equation is written on the assumption that before the collision one of the particles was at rest. In the center-of-mass system (31.31) is the same as the cross section for radiation in the field of the nucleus (with $Z = 1$) in the nonrelativistic approximation [see (31.15)].

§ 32. Emission of Long-Wavelength Photons

1. Infrared "Catastrophe"

In the previous section we saw that the matrix element which gives the radiation by an electron in an external field approaches infinity as $\omega^{-3/2}$ when the photon energy ω approaches zero. Therefore the probability that an electron will emit a photon whose energy lies between ω and $\omega + d\omega$ is proportional, for low photon energies, to

1) G. Garibyan, Bull. Acad. Sci. Armenian SSR 5, 3 (1952).

$$d\omega \sim \omega^{-2} d\omega = \frac{d\omega}{\omega}, \quad (32.1)$$

and the total emission probability diverges logarithmically as $\omega \rightarrow 0$. This divergence in the low-energy region is called the infrared catastrophe. It should be noted, however, that this situation has nothing in common with the fundamental divergences of quantum electrodynamics in the high-momentum regions of the virtual particles, and is related to the fact that ordinary perturbation theory based on the series expansion of the S matrix in powers of e is not valid for processes involving long-wavelength photons. Indeed by repeating the considerations which led to (32.1), it is easy to show that if the probability $\mathcal{P}_1$ for the emission of a single long-wavelength photon is proportional to $e^2 \ln \frac{c}{\omega}$, where c is an energy of the same order of magnitude as that of the electron, then the probability $\mathcal{P}_2$ for the emission of two photons is proportional to $(e^2 \ln \frac{c}{\omega})^2$. Therefore, the order of magnitude of the ratio of the probabilities is given by

$$\xi = \frac{\mathcal{P}_2}{\mathcal{P}_1} \sim e^2 \ln \frac{c}{\omega}, \quad \omega \rightarrow 0, \quad (32.2)$$

It is this ratio, and not the quantity e^2 , as we have thus far assumed, which is the power series parameter in the application of perturbation theory to processes involving the interaction of the electron with long-wavelength photons. Strictly speaking, since ξ is not small compared to unity when $\omega \rightarrow 0$, perturbation theory is not applicable to these cases.

The inapplicability of ordinary perturbation theory is related to the fact that the number of photons emitted by an electron per unit energy interval approaches infinity as $\omega \rightarrow 0$, whereas perturbation theory assumes that the radiation of a single photon is always more probable than that of two or more.

In order to show that as $\omega \rightarrow 0$ the number of photons emitted actually approaches infinity, we note that if the energy and momentum of the photon are much smaller than the kinetic energy and momentum change of the electron, and if the photon wavelength is much larger than the classical electron radius, then we may consider the electron motion to be given and may use classical electrodynamics. Assuming for simplicity that the velocity of the electron is small with respect to that of light, we can make use of the following formula for the intensity $\mathcal{E}_\omega$ of the dipole radiation in a frequency interval $d\omega$:

$$d\mathcal{E}_\omega = \frac{8\pi}{3c^3} |\ddot{d}_\omega|^2 d\omega, \quad (32.3)$$

where $\ddot{d}_\omega$ is the Fourier component of the second time derivative of the dipole moment,

$$\ddot{d}_\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \ddot{d}(t) e^{i\omega t} dt.$$

If $\omega \rightarrow 0$, then

$$\ddot{d}_{\omega \rightarrow 0} = \frac{1}{2\pi} (\ddot{d}_2 - \ddot{d}_1)$$

¹⁾ See, for instance, L. Landau and E. Lifshits, Field Theory (State Tech. Press, 1948).

(d_1 and d_2 are the values of the dipole moment before and after radiation). In the case in which we are interested, $\ddot{d} = \ddot{xy}$ and $\ddot{d}_{\omega \rightarrow 0} = \frac{c}{2} (\ddot{y}_2 - \ddot{y}_1)$ where y_1 and y_2 are the electron velocities before and after radiation. Therefore

$$d\mathcal{E}_{\omega \rightarrow 0} = \frac{2}{3\pi} \frac{c^2}{v^2} (\ddot{y}_2 - \ddot{y}_1)^2 d\omega. \quad (32.4)$$

We see that the intensity of the radiation per unit frequency interval $\frac{d\mathcal{E}}{d\omega}$ has a finite nonzero limit as $\omega \rightarrow 0$. It follows from this that the mean number of photons emitted, namely $\frac{1}{\hbar\omega} \frac{d\mathcal{E}_\omega}{d\omega}$, approaches infinity as $\omega \rightarrow 0$, as was asserted above.

Since the probability for an electron transition from a state with momentum p_1 to one with momentum p_2 is always finite, the probability of simultaneous emission of an infinite number of photons with infinitesimally small frequencies ($\omega \rightarrow 0$) is also finite and nonzero. Therefore the probability for radiating one or a finite number of photons as $\omega \rightarrow 0$ actually vanishes, and does not become infinite as is assumed by perturbation theory.

The quantity $\frac{1}{\hbar\omega} \frac{d\mathcal{E}_\omega}{d\omega}$ is the average number of photons with frequency ω emitted by an electron into the frequency interval $d\omega$. We shall now find the probability that an electron emits some arbitrary number n of long-wavelength photons whose frequencies lie in the interval $\omega_1 \leq \omega \leq \omega_2$. Assuming, as before, that

$$\frac{\hbar\omega}{\epsilon} \ll 1, \quad \frac{\hbar\omega}{c\Delta p} \ll 1, \quad \frac{e^2\omega}{mc^2} = \frac{r_0}{\lambda} \ll 1, \quad (32.5)$$

where ϵ is the kinetic energy of the electron, Δp is its momentum change λ is the photon wavelength, and r_0 is the classical electron radius, we may assume that the emission of photons does not effect the motion of the electron, that is, we may consider its motion given. Under these conditions successive photon emission events are statistically independent, and therefore the probability for photon emission is given by Poisson's formula

$$w_n = W^n \frac{e^{-W}}{n!}, \quad (32.6)$$

where $W = \bar{n}$ is the average number of emitted photons whose frequencies lie in the given interval $\omega_1 \leq \omega \leq \omega_2$.

When conditions (32.5) are satisfied W can be found from classical electrodynamics. In particular, if $d\mathcal{E}_\omega = I_\omega d\omega d\Omega$ is the classical radiation intensity in the frequency interval $(\omega, \omega + d\omega)$ and the solid angle $d\Omega$, then

$$W = \bar{n} = \int_{\omega_1}^{\omega_2} \int_{\Omega} \frac{I_\omega}{\hbar\omega} d\omega d\Omega. \quad (32.7)$$

²⁾ Here $\hat{n}$ is the unit vector in the direction of radiation.

We shall now show how to find $\mathcal{J}_{\omega\Omega}$. It is known that

$$d\mathcal{J}_{\omega\Omega} = c |H_{\omega}|^2 R^2 d\omega d\Omega,$$

where H_{ω} is the Fourier component corresponding to the frequency ω of the magnetic field $\mathbf{H}$ at the point $\mathbf{r}$; the quantity H_{ω} is related to the Fourier component of the potential A_{ω} by the expression $H_{\omega} = \frac{1}{c} (\mathbf{k} A_{\omega})$. Using the Liénard-Wiechert potential¹⁾, it is easy to show that

$$A_{\omega} = \frac{e}{2\pi c R} \int_{-\infty}^{\infty} \mathbf{v}(t) e^{i(\omega t - kr(t))} dt, \quad \varphi_{\omega} = \frac{e}{2\pi c R} \int_{-\infty}^{\infty} e^{i(\omega t - kr(t))} dt, \quad (32.8)$$

where $\mathbf{r} = \mathbf{r}(t)$ is the equation of the electron trajectory, and $\mathbf{v}(t)$ is the electron velocity (φ_{ω} is the Fourier component of the scalar potential). Therefore

$$d\mathcal{J}_{\omega\Omega} = c |H_{\omega}|^2 R^2 d\omega d\Omega = c R^2 k^2 (|A_{\omega}|^2 - |\varphi_{\omega}|^2) d\omega d\Omega = \\ = \frac{e^2 c}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{R^2} \mathbf{v}(t) \mathbf{v}(t') - 1 \right) e^{i(\omega(t-t') - k(r(t) - r(t')))} dt dt' k^2 d\omega d\Omega \quad (32.9)$$

(we have made use of the relation $\frac{k A_{\omega}}{c} + \frac{1}{c} \omega \varphi_{\omega} = 0$).

If the frequency ω satisfies the condition

$$\omega \tau \ll 1,$$

where τ is of the same order of magnitude as the time during which the electron scattering takes place, then on integrating (32.9) we may consider that

$$\begin{aligned} r(t) &\approx \mathbf{v}_1 t + \mathbf{a}, & -\infty < t < 0, \\ r(t) &\approx \mathbf{v}_2 t + \mathbf{b}, & 0 \leq t < \infty \end{aligned} \quad (32.10)$$

($\mathbf{v}_1$ and $\mathbf{v}_2$ are the electron velocities before and after radiation, and $\mathbf{a}$ and $\mathbf{b}$ are constants). Inserting (32.10) into (32.9), we obtain

$$d\mathcal{J}_{\omega\Omega} = \frac{e^2}{4\pi^2 c^3} \left(\left(\frac{\mathbf{v}_2}{1 - n v_2/c} - \frac{\mathbf{v}_1}{1 - n v_1/c} \right)^2 - \left(\frac{n v_2}{1 - n v_2/c} - \frac{n v_1}{1 - n v_1/c} \right)^2 \right) d\omega d\Omega. \quad (32.11)$$

¹⁾ See the note on p. 330.

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are much smaller than the velocity of light, then (32.11) reduces to (32.4)¹⁾.

Let us now determine the probability for electron scattering in an external field accompanied by radiation of n long-wavelength photons. Assuming that conditions (32.5) are satisfied, which means that the recoil effect of the radiation on the electron motion is extremely small, we can write the probability for electron scattering with radiation of n long-wavelength photons in the form

$$dw_{en} = w_n dw_e, \quad (32.12)$$

where $dw_e = f d\Omega_e$ is the probability for elastic scattering of the electron into a solid angle $d\Omega_e$, as defined in Section 13. Clearly the frequency interval (ω_1, ω_2) we must use should contain the frequency zero. The average number of photons emitted in such an interval is infinite according to (32.7), and therefore $w_n = 0$. In other words, the probability for electron scattering with the emission of a finite number of long-wavelength photons is zero. In particular, the probability for pure elastic scattering vanishes. Since, on the other hand,

$$\sum_{n=0}^{\infty} w_n = 1,$$

we have

$$\sum_{n=0}^{\infty} dw_{en} = dw_e. \quad (32.13)$$

It may therefore be said that the probability for elastic scattering of an electron dw_e , as found in Section 13, which does not take into account the interaction between the electron and the radiation field, is the total scattering probability for the electron independent of the number of long-wavelength photons it emits. This probability, as we see, can be determined in the case of a sufficiently weak external field by using perturbation theory, which is not possible for the probability for scattering accompanied by the emission of a finite number of photons as $\omega \rightarrow 0$; according to perturbation theory this probability becomes infinite, whereas in actuality it vanishes.

The mean energy of the long-wavelength photons emitted in the energy interval ($\omega, \omega + d\omega$) and into the solid angle $d\Omega$ when an electron is scattered into the solid angle $d\Omega_e$ can, according to (32.6), be written

$$\overline{d\epsilon} = \sum_{n=0}^{\infty} n \omega e^{-W} \frac{W^n}{n!} d\omega_n = \omega W d\omega_n.$$

We note that in this case $W = \frac{d\mathcal{J}_{\omega\Omega}}{\omega}$, so that we finally obtain

$$\overline{d\epsilon} = d\mathcal{J}_{\omega\Omega} d\omega_n. \quad (32.14)$$

¹⁾ Equation (32.11), like the other equations of classical electrodynamics [(32.3), (32.4), (32.8), (32.9)], is here written in ordinary (not Heaviside) units.

This quantity, as may have been expected, is the product of the total probability dW for electron scattering into a solid angle $d\Omega$, by the classical expression dW_{cl} for the energy radiated in this process. The same result is obtained from ordinary perturbation theory in the low-frequency region. We thus see that the average energy radiated and the total scattering probability for the electron (independent of the number of long-wavelength photons emitted) can be determined with the aid of ordinary perturbation theory both for high and low frequencies, although the probability for scattering of an electron with the emission of a finite number of photons as $\omega \rightarrow 0$ is given incorrectly by perturbation theory.

2. Investigation of the Divergence in the Low-Frequency Region.

We shall show that the Poisson distribution of (32.6) and expression (32.7) for the average number of photons emitted can both be simply obtained from the general equations of quantum electrodynamics¹⁾.

When conditions (32.5) are fulfilled, the current density can be treated as a given function of space and time, since what these conditions imply is that the radiation hardly effects the motion of the electron which we may therefore consider given. In this case we can use the following expression for the S matrix

$$S = e^{-i \int_{-\infty}^{\infty} j_{\mu} A_{\mu} dx} \quad (32.15)$$

where A_{μ} is the electromagnetic field operator, and j_{μ} is the current density, but treated as a c-number rather than an operator.

Expression (32.15) differs from the usual expression for the S matrix in that the chronological operator is absent. If j_{μ} is considered a c-number, the expressions differ only by a phase factor

$$e^{\frac{i}{2} \int j_{\mu}(x) D(x-x') j_{\mu}(x') dx dx'}$$

of modulus unity. This phase factor may be dropped, since it does not enter into the expression for the probability of photon emission.

Our problem consists of finding the matrix element, or more exactly that part of the S matrix, which corresponds to the emission of some definite number n of photons. It is clear that we must take account not only of the n -th term in the expansion of (32.15) in a power series in its exponent, but also of terms of higher order, since terms of order $n+2p$ (where p is a positive integer) will also give nonzero matrix elements for the process of interest. Indeed, these terms can describe emission of n real, and p virtual photons, followed by the absorption of the virtual ones.

Let $\mathcal{H}$ be some Hermitian operator which, like $A_{\mu}(x)$ contains, a sum of photon emission and absorption operators, and let us consider the operator $\mathcal{H}^m$, where $m \geq n$. Since of the m factors $\mathcal{H}$ we can choose n of them in $\binom{m}{n}$ ways, it is easy to see that the matrix element, or more exactly that part of the operator $\mathcal{H}^m$, corresponding to emission of n photons can be written

$$(\mathcal{H}^m)_n = \binom{m}{n} (\mathcal{H}^n)_n (\mathcal{H}^{m-n})_0, \quad (32.16)$$

where $(\mathcal{H}^{m-n})_0$ is the vacuum expectation value of $\mathcal{H}^{m-n}$. This expression can be written symbolically in the form

$$(\mathcal{H}^m)_n = \frac{1}{n!} (\mathcal{H}^m)_n \left(\frac{d^n}{d\lambda^n} \mathcal{H}^m \right)_0,$$

from which it follows that any function $f(\mathcal{H})$, which can be written as a power series satisfies the relation

$$(f(\mathcal{H}))_n = \frac{1}{n!} (\mathcal{H}^n)_n (f(\mathcal{H}))_0. \quad (32.17)$$

In particular, that part of S which corresponds to emission of n photons can be written

$$(S)_n = \frac{1}{n!} \left(i \int j_{\mu} A_{\mu} dx \right)^n (e^{i \int j_{\mu} A_{\mu} dx})_0. \quad (32.18)$$

We shall now show how to calculate the vacuum expectation value of the S matrix. To do this, let us treat the quantity $(e^{i\mathcal{H}})_0$ as a function of the parameter λ . The derivative of this function with respect to λ is

$$\frac{d}{d\lambda} (e^{i\mathcal{H}})_0 = (i\mathcal{H} e^{i\mathcal{H}})_0 = (i\mathcal{H})_0 (e^{i\mathcal{H}})_0. \quad (32.19)$$

It is clear that part of the operator $e^{i\mathcal{H}}$, which gives a nonzero contribution to (32.19) must correspond to emission of a single photon. According to (32.17), this part is $\lambda \mathcal{H} (e^{i\mathcal{H}})_0$; therefore,

$$\frac{d}{d\lambda} (e^{i\mathcal{H}})_0 = \lambda (e^{i\mathcal{H}})_0 (e^{i\mathcal{H}})_0.$$

The solution of this equation which vanishes when $\lambda = 0$ is of the form

$$(e^{i\mathcal{H}})_0 = e^{\frac{1}{2} \lambda^2 (\mathcal{H}^2)_0}. \quad (32.20)$$

Setting $\mathcal{H} = \int j_{\mu} A_{\mu} dx$ and $\lambda = i$, we finally obtain

$$(e^{i \int j_{\mu} A_{\mu} dx})_0 = e^{-\frac{1}{2} W}, \quad (32.21)$$

¹⁾ R. Glauber Phys. Rev. 84, 395 (1951).

where

$$W = \int J_p(x) \langle A_p(x) A_p(x') \rangle_0 J_p(x') dx dx' = \frac{1}{2} \int J_p(x) D^{(1)}(x-x') J_p(x') dx dx' \quad (32.22)$$

and $D^{(1)}(x)$ is given by (16.22), namely

$$D^{(1)}(x) = \frac{1}{(2\pi)^3} \int e^{ikx} \tilde{D}(k^2) d^4k.$$

Having determined the form of $(S)_n$, let us now find the probability for the emission of n photons. This probability, which we shall denote by w_n , is clearly

$$w_n = \langle \Phi(+\infty), \Phi(+\infty) \rangle,$$

where

$$\Phi(+\infty) = (S)_n \Phi(-\infty).$$

Inserting (32.18) for $(S)_n$ into this expression, we obtain

$$w_n = \frac{1}{(n!)^2} \left(\left(\int J_p A_p dx \right)^n; \left(\int J_p A_p dx \right)^n \right)_0 e^{-W}. \quad (32.23)$$

In this expression the semicolon breaks up the expectation value into two factors, the one on the right causing emission of n photons, and the one on the left causing their absorption. Since the A_p operators can be divided into pairs of one operator from each factor in $n!$ ways, and since each such way of forming pairs leads to a factor W , we finally obtain

$$w_n = \frac{W^n}{n!} e^{-W}.$$

Thus, the probability of emitting n photons is given by the Poisson distribution. We are left with the problem of showing that W as defined by (32.23) is the same as that defined by (32.7). To do this let us replace $\int_\mu$ by

$$J_p = (ev\delta(r-r(t)), i\epsilon\delta(r-r(t))).$$

Using the definition (10.22) of $D^{(1)}(x)$ and eliminating the δ -function by integrating over dr and dr' , we obtain (32.7), where $d\mathcal{F}_{\omega n}$ is given by (32.9).

3-The Bloch-Nordsieck Method.

The interaction of an electron with long-wavelength photons can also be treated by another method¹⁾. We introduce the Hamiltonian for the electron and system of photons

$$H = (\alpha, p - eA) + \beta m + \frac{1}{2} \int (E^2 + H^2) dV, \quad (32.24)$$

where the vector potential is given by the expansion

$$A = \frac{1}{\sqrt{V}} \sum_{\lambda} e_{\lambda}^{-1} e_{\lambda} (P_{\lambda} \cos k_{\lambda} r + Q_{\lambda} \sin k_{\lambda} r); \quad (32.25)$$

here the index λ is used to label the wave vector k_{λ} and frequency ω_{λ} of the photon e_{λ} is a unit vector in the direction of polarization, whose direction is labeled by the index λ , and V is the normalizing volume. The dynamical variables P_{λ} and Q_{λ} satisfy the commutation relations

$$[P_{\lambda}, Q_{\lambda'}] = -i\delta_{\lambda\lambda'}, \quad [P_{\lambda}, P_{\lambda'}] = 0, \quad [Q_{\lambda}, Q_{\lambda'}] = 0.$$

Using (32.25) to calculate the electric and magnetic fields E and H , we rewrite H in the form

$$H = (\alpha, p - \sum_{\lambda} e_{\lambda} [P_{\lambda} \cos k_{\lambda} r + Q_{\lambda} \sin k_{\lambda} r]) + \beta m + \frac{1}{2} \sum_{\lambda} \omega_{\lambda} (P_{\lambda}^2 + Q_{\lambda}^2), \quad (32.26)$$

where

$$e_{\lambda} = e(V\omega_{\lambda})^{-1/2} e_{\lambda}.$$

Let us now consider the Schroedinger equation

$$H\psi = E\psi \quad (32.27)$$

and separate ψ into two parts ψ^+ and ψ^- where

$$\psi = \psi^+ + \psi^-,$$

¹⁾ F. Bloch and A. Nordsieck, Phys. Rev. 5, 54 (1937).

which satisfy the equations

$$\Delta\psi^+ = \psi^+, \quad \Delta\psi^- = -\psi^-$$

where

$$\Delta = \alpha v + \beta \sqrt{1-v^2}, \quad v = \text{const.}$$

Using the relation

$$\Delta H\psi = \mathcal{E}\Delta\psi, \quad (32.27)$$

and adding and subtracting Equations (32.27) and (32.27'), we obtain

$$\begin{aligned} \left\{ \left(v, p - \sum_{\mathbf{a}} a_{\mathbf{a}} [P_{\mathbf{a}} \cos k_{\mathbf{a}} r + Q_{\mathbf{a}} \sin k_{\mathbf{a}} r] \right) + m(1-v^2)^{1/2} \pm \right. \\ \left. \pm \frac{1}{2} \sum_{\mathbf{a}} \omega_{\mathbf{a}} (P_{\mathbf{a}}^2 + Q_{\mathbf{a}}^2) - \mathcal{E} \right\} \psi^{\pm} = - \left\{ (v \pm \alpha) p - m(1-v^2)^{-1/2} v - \right. \\ \left. - \sum_{\mathbf{a}} a_{\mathbf{a}} [P_{\mathbf{a}} \cos k_{\mathbf{a}} r + Q_{\mathbf{a}} \sin k_{\mathbf{a}} r] \right\} \psi^{\mp}. \end{aligned} \quad (32.28)$$

We now note that if the interaction between the electron and the field is allowed to approach zero, the functions ψ^+ and ψ^- will describe electron states with energies $\frac{m}{\sqrt{1-v^2}}$ and $-\frac{m}{\sqrt{1-v^2}}$. Therefore, in order to obtain positive energy states, we must set ψ^- equal to zero on the right side of (32.28); this means that we are approximating the wave function by setting $\psi^+ = \psi$, obtaining

$$\begin{aligned} \left\{ \left(v, p - \sum_{\mathbf{a}} a_{\mathbf{a}} [P_{\mathbf{a}} \cos k_{\mathbf{a}} r + Q_{\mathbf{a}} \sin k_{\mathbf{a}} r] \right) + m\sqrt{1-v^2} + \right. \\ \left. + \frac{1}{2} \sum_{\mathbf{a}} (P_{\mathbf{a}}^2 + Q_{\mathbf{a}}^2) \omega_{\mathbf{a}} - \mathcal{E} \right\} \psi = 0. \end{aligned} \quad (32.29)$$

We shall later see what error is involved in this approximation.

In order to solve (32.29), which contains no Dirac matrices, we shall apply a canonical transformation to the Hamiltonian

$$\begin{aligned} H = \left(v, p - \sum_{\mathbf{a}} a_{\mathbf{a}} [P_{\mathbf{a}} \cos k_{\mathbf{a}} r + Q_{\mathbf{a}} \sin k_{\mathbf{a}} r] \right) + m\sqrt{1-v^2} + \\ + \frac{1}{2} \sum_{\mathbf{a}} \omega_{\mathbf{a}} (P_{\mathbf{a}}^2 + Q_{\mathbf{a}}^2) \end{aligned} \quad (32.29')$$

to eliminate those terms which are linear in $P_{\mathbf{a}}$ and $Q_{\mathbf{a}}$.

We recall that in general a canonical transformation is performed in the following way. Let p_{α}, x_{α} be canonically-conjugate variables satisfying the condition $[p_{\alpha}, x_{\alpha}] = -i$, where p_{α} is the derivative with respect to x_{α} , namely $p_{\alpha} = -i \frac{\partial}{\partial x_{\alpha}}$. If we transform to new canonical variables p'_{α}, x'_{α} which must still

satisfy the condition $[p'_{\alpha}, x'_{\alpha}] = -i$, then in general p'_{α} will not be the derivative with respect to x'_{α} . It is possible, however, to transform the wave function so as to have the operator p'_{α} act on the new wave function ψ' as the derivative with respect to x'_{α} . Indeed, performing the transformation $p_{\alpha} \rightarrow p'_{\alpha}, x_{\alpha} \rightarrow x'_{\alpha}$ in the Hamiltonian $H(p_{\alpha}, x_{\alpha})$, we obtain the Hamiltonian $H'(p'_{\alpha}, x'_{\alpha}) = H(p_{\alpha}, x_{\alpha})$. We now define the function $S(x'_{\alpha})$ such that

$$H'(p'_{\alpha}, x'_{\alpha}) = SH' \left(-i \frac{\partial}{\partial x'_{\alpha}}, x'_{\alpha} \right) S^{-1}. \quad (32.30)$$

Inserting this expression into the equation for the wave function ψ

$$H(p_{\alpha}, x_{\alpha})\psi = H'(p'_{\alpha}, x'_{\alpha})\psi = \mathcal{E}\psi,$$

we obtain

$$SH' \left(-i \frac{\partial}{\partial x'_{\alpha}}, x'_{\alpha} \right) S^{-1} \psi = \mathcal{E}\psi,$$

whence

$$H' \left(-i \frac{\partial}{\partial x'_{\alpha}}, x'_{\alpha} \right) \psi' = \mathcal{E}\psi', \quad (32.31)$$

where

$$\psi' = S^{-1} \psi. \quad (32.32)$$

We thus see that when performing a canonical transformation $x_{\alpha} \rightarrow x'_{\alpha}, p_{\alpha} \rightarrow p'_{\alpha}$, we must also transform the wave function according to (32.32), where S is given by (32.30).

In our case, the Hamiltonian is given by (32.29'). We shall perform the canonical transformation given by

$$\begin{aligned} P_{\mathbf{a}} &= P'_{\mathbf{a}} + a_{\mathbf{a}} \cos k_{\mathbf{a}} r, \\ Q_{\mathbf{a}} &= Q'_{\mathbf{a}} + a_{\mathbf{a}} \sin k_{\mathbf{a}} r, \quad r = r', \\ p &= p' - \sum_{\mathbf{a}} k_{\mathbf{a}} a_{\mathbf{a}} \left[P'_{\mathbf{a}} \cos k_{\mathbf{a}} r + Q'_{\mathbf{a}} \sin k_{\mathbf{a}} r + \frac{1}{2} \omega_{\mathbf{a}} \right] \end{aligned} \quad (32.33)$$

and choose the constants σ_{λ} so that the transformed Hamiltonian $\mathcal{H}'$ does not contain terms linear in $\mathbf{p}'_{\lambda}$ and $\mathbf{Q}'_{\lambda}$. It is easy to see that this condition leads to the following values for σ_{λ} :

$$\sigma_{\lambda} = \frac{\sigma_{\lambda 0}}{|\mathbf{k}_{\lambda}| - v \mathbf{k}_{\lambda}}. \quad (32.34)$$

Putting $S = e^{-f}$ in (32.30), we obtain the following equation for f :

$$\begin{aligned} \frac{\partial f}{\partial Q_{\lambda}} &= -i \sigma_{\lambda} \cos \mathbf{k}_{\lambda} \mathbf{r} - i v \frac{\partial f}{\partial \mathbf{r}} - \frac{1}{2} \sum_{\lambda} \omega_{\lambda} \left[\frac{\partial^2 f}{\partial Q_{\lambda}^2} + \left(\frac{\partial f}{\partial Q_{\lambda}} \right)^2 \right] = \\ &= - \sum_{\lambda} v \sigma_{\lambda} \left(Q_{\lambda} \sin \mathbf{k}_{\lambda} \mathbf{r} + \sigma_{\lambda} \sin^2 \mathbf{k}_{\lambda} \mathbf{r} - \frac{1}{2} \sigma_{\lambda} \right) + \\ &\quad + \sum_{\lambda} \left(Q_{\lambda}^2 \omega_{\lambda} \sin \mathbf{k}_{\lambda} \mathbf{r} + \frac{1}{2} \omega_{\lambda} \sigma_{\lambda}^2 \sin^2 \mathbf{k}_{\lambda} \mathbf{r} \right), \end{aligned}$$

whose solution is

$$f = -i \sum_{\lambda} \sigma_{\lambda} \cos \mathbf{k}_{\lambda} \mathbf{r} \left[Q_{\lambda} + \frac{1}{2} \sigma_{\lambda} \sin \mathbf{k}_{\lambda} \mathbf{r} \right], \quad S = e^{-f}. \quad (32.35)$$

The equation for the transformed wave function $\psi' = S^{-1} \psi$ is

$$\left(v \mathbf{p}' + \frac{m}{\sqrt{1-v^2}} + \frac{1}{2} \sum_{\lambda} (P_{\lambda}^2 + Q_{\lambda}^2) \omega_{\lambda} - \frac{1}{2} \sum_{\lambda} \frac{(v \sigma_{\lambda})^2}{|\mathbf{k}_{\lambda}| - v \mathbf{k}_{\lambda}} - \mathcal{E} \right) \psi' = 0 \quad (32.36)$$

and has the following solution:

$$\left. \begin{aligned} \psi' &= v(\mathbf{v}) V^{-1/2} e^{i m(1-v^2)^{-1/2} \mathbf{v} \cdot \mathbf{r}} \prod_{\lambda} h_{n_{\lambda}}(Q_{\lambda}), \\ \mathcal{E} &= \frac{m}{\sqrt{1-v^2}} + \sum_{\lambda} \left(n_{\lambda} + \frac{1}{2} \right) \omega_{\lambda} - \frac{1}{2} \sum_{\lambda} \frac{(v \sigma_{\lambda})^2}{|\mathbf{k}_{\lambda}| - v \mathbf{k}_{\lambda}}, \end{aligned} \right\} \quad (32.37)$$

where $h_n(x)$ is the normalized solution of the oscillator equation

$$h_n''(x) - x^2 h_n(x) + (2n+1) h_n(x) = 0$$

and $v(\mathbf{v})$ is a spinor satisfying the relation $\Lambda v = v$.

Transforming back to the original variables, we can write the approximate wave function of the form

$$\begin{aligned} \psi &= \frac{v(\mathbf{v})}{\sqrt{V}} \exp \left\{ i m(1-v^2)^{-1/2} \mathbf{v} \cdot \mathbf{r} + i \sum_{\lambda} \sigma_{\lambda} \cos \mathbf{k}_{\lambda} \mathbf{r} \left[Q_{\lambda} + \frac{1}{2} \sigma_{\lambda} \sin \mathbf{k}_{\lambda} \mathbf{r} \right] \right\} \times \\ &\quad \times \prod_{\lambda} h_{n_{\lambda}}(Q_{\lambda} + \sigma_{\lambda} \sin \mathbf{k}_{\lambda} \mathbf{r}). \end{aligned} \quad (32.38)$$

In order to find the error involved in replacing ψ by ψ' , let us insert (32.38) into the right side of Equation (32.28). We then obtain an expression for ψ'' . The solution of this equation shows that $|\psi''|^2/|\psi|^2$ is of order of magnitude

$$\frac{|\psi''|^2}{|\psi|^2} \sim \frac{e^2 \omega_m}{m} \left[f(v) \frac{\omega_m}{m} + g(v) \frac{e^2 \omega_m}{m} \right], \quad (32.39)$$

where ω_m is the highest frequency included in our considerations, and $f(v)$ and $g(v)$ are certain functions of v . If $v \ll 1$, then $f \sim \frac{1}{2\pi}$ and $g \sim \frac{4}{9\pi} v^2$; if $v \sim 1$, then $f \sim \frac{1}{4\pi} (1-v^2)$, $g \sim \frac{1}{4\pi} (1-v^2)^2$. Equation (32.39) shows that this method for treating Equation (32.27) is essentially a power series not in e^2 , but in $\frac{e^2 \omega}{m}$ and $\frac{\omega}{m}$.

It can be shown that the use of ordinary perturbation theory with the wave functions of (32.38) leads to expression (32.12) for the probability of scattering an electron in a weak external field with the emission of n photons, if we neglect the momentum of the photon in comparison with the change of momentum of the electron.

Thus we have shown that perturbation theory, which leads to incorrect results for the probability of radiation at low frequencies, nevertheless gives the correct value for the total scattering probability as well as for the mean energy radiated in this frequency region.

The probability for radiating photons whose energy is not extremely small can be determined on the basis of ordinary perturbation theory, that is, by expanding the matrix elements in a power series in e ; the probability for radiating a single photon, however, should then be interpreted not as the probability that only this one photon is radiated, but as the probability that in addition to this photon an arbitrary number of long-wavelength photons whose frequencies $\omega \rightarrow 0$ are also radiated.

4. Relation Between the Photon "Mass" and the Minimum Frequency.

Since perturbation theory gives incorrect values for the probabilities of various processes involving the interaction of an electron with long-wavelength photons, henceforth in using perturbation theory we shall separate out the low-frequency region; i.e., we shall assume that the photon momentum is greater than some minimum value $\omega_{\min}$. This value can be excluded by special investigations of the interaction of the electron with long-wavelength photons (see, for instance, problems involving radiative corrections to electron scattering and the radiative atomic level shift in Sections 45 and 46).

In practice it is more convenient, however, to make use of another condition, which is equivalent to $\omega > \omega_{\min}$: this does not involve restricting the frequency of the photon, but involves assuming that the photon has a certain extremely small nonzero mass λ^3 . The introduction of this mass, which, as opposed to $\omega_{\min}$, is relativistically invariant, brings great simplification to the calculations.

We shall now show how to find the relation between the photon "mass" λ and the minimum frequency

³⁾ Feynman, Phys. Rev. 76, 769 (1949).

$\omega_{\min}$. We shall start with expression (31.4) for the matrix element $S_{1 \rightarrow 1}^{(0)}$ which determines the radiation of a photon with energy $\hbar\omega$ and polarization in the direction μ :

$$S_{1 \rightarrow 1}^{(0)} = -ie\vec{u}_2 \left(\frac{\gamma_p}{\sqrt{2k_0}} \frac{i(\hat{p}_2 + \hat{k}) - m}{(p_2 + k)^2 + m^2} \hat{a}(q) + \hat{a}(q) \frac{i(\hat{p}_1 - \hat{k}) - m}{(p_1 - k)^2 + m^2} \frac{\gamma_p}{\sqrt{2k_0}} \right) u_1, \quad (32.40)$$

$$q = p_2 - p_1 + k.$$

Since we are interested in the low-frequency region, we shall drop $\hat{k}$ from the numerators of the fractions in (32.40). Introducing the photon "mass" λ according to the relation

$$k^0 + \lambda^2 = 0,$$

let us rewrite (32.40) in the form

$$S_{1 \rightarrow 1}^{(0)} = \frac{ie^2}{\sqrt{2k_0}} \vec{u}_2 \left(\frac{i\hat{p}_2 - m}{\gamma_p \lambda^2 - 2p_2 k} \hat{a}(q) + \hat{a}(q) \frac{i\hat{p}_1 - m}{\lambda^2 + 2p_1 k} \gamma_p \right) u_1, \quad (32.40')$$

where

$$pk = p_k - ek_0, \quad e = \sqrt{p^2 + m^2}, \quad k_0 = \sqrt{k^2 + \lambda^2}.$$

Since $(\hat{p}_1 + \underline{m})u_1 = 0$, we have

$$\vec{u}_2 \hat{a}(q) (i\hat{p}_1 - m) u_1 = i\vec{u}_2 \hat{a}(q) (\hat{p}_1 \gamma_p + \gamma_p \hat{p}_1) u_1 = 2i\vec{u}_2 \hat{a}(q) u_1 p_{1p}.$$

Similarly

$$\vec{u}_2 \gamma_p (\hat{p}_2 - m) \hat{a}(q) u_1 = 2i\vec{u}_2 \hat{a}(q) u_1 p_{2p}.$$

Therefore

$$S_{1 \rightarrow 1}^{(0)} = \frac{e^2}{\sqrt{2k_0}} (\vec{u}_2 \hat{a}(q) u_1) \left(\frac{p_{2p}}{-\lambda^2/2 + p_{2k}} + \frac{p_{1p}}{-\lambda^2/2 - p_{1k}} \right). \quad (32.41)$$

The term λ^2 can be neglected in the denominators here, since p_k contains the term $\epsilon \sqrt{k^2 + \lambda^2}$ which is much greater than λ^2 (we assume that the electron energy $\epsilon \gg \lambda$).

Let us now determine the cross section for scattering of an electron by an external field with the emission of a photon whose energy is no greater than $\Delta\epsilon$. We shall assume that $\Delta\epsilon$ is much smaller than the electron energy ϵ and much greater than the photon mass λ :

$$\lambda \ll \Delta\epsilon \ll \epsilon.$$

The differential cross section for this scattering process is (see Section 28)

$$d\sigma = \frac{2\pi e^4}{v} \sum_{\sigma_1, \sigma_2} |\vec{u}_2 \hat{a}_\sigma u_1|^2 \frac{B}{(2\pi)^3} p_f d\sigma, \quad \hat{a}_\sigma = \int \hat{A}^{(\sigma)}(x) e^{-i\sigma x} d^3x, \quad (32.42)$$

where

$$B = \sum_{p=1}^4 \int_{\Delta\epsilon} \left(\frac{p_{2p}}{p_k k} - \frac{p_{1p}}{p_k k} \frac{d^3k}{k_0} \right), \quad (32.43)$$

$d\sigma$ is the element of solid angle about the momentum p_2 of the electron after the scattering, v is the electron velocity, which is equal to the electron current density (the normalizing volume is assumed unity), and

$p_f = \frac{p_f^2 d^3p}{(2\pi)^3 d\epsilon} = \frac{|p_2| \epsilon}{(2\pi)^3}$ is the number of final electron states per unit energy interval and unit

solid angle; a summation over spins of the initial and final electron states is included in (32.42).

We note that for the Coulomb field of the nucleus

$$\hat{a}_\sigma = -i\beta \frac{Zv}{|q|^2}$$

and

$$\sum_{\sigma_1, \sigma_2} |\vec{u}_2 \hat{a}_\sigma u_1|^2 = 2 \left| \frac{Ze}{q^2} \right|^2 \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) = 8\pi Z^2 \frac{v^2}{q^4} \left(1 - v^2 \sin^2 \frac{\theta}{2} \right), \quad (32.44)$$

where θ is the scattering angle, $-Ze$ is the charge of the nucleus, and α is the fine structure constant.

In order to establish the relation between the photon mass λ and the minimum frequency $\omega_{\min}$, we shall calculate the quantity B which enters into (32.42) in two independent ways: first we shall assume that $\lambda \neq 0$, and that k_0 varies from $k_0 = \lambda$ to $k_0 = \Delta\epsilon$; then we shall assume that $\lambda = 0$, and k_0 varies from $k_0 = \omega_{\min}$ to $k_0 = \Delta\epsilon$. By comparing the expressions so obtained for B , which we shall denote by B_λ and $B_{\omega_{\min}}$ respectively we can find the relation between λ and $\omega_{\min}$.

Noting that $p_{1,2}^2 + m^2 = 0$, Equation (32.43) gives

$$B_\lambda = - \int_{k=0}^{\Delta\epsilon} \frac{k^2 d|k| d\sigma_k}{\sqrt{k^2 + \lambda^2}} \left\{ \frac{m^2}{(\epsilon_2 \sqrt{k^2 + \lambda^2} - p_k k)^2} + \frac{m^2}{(\epsilon_1 \sqrt{k^2 + \lambda^2} - p_k k)^2} + \frac{2p_1 p_2}{(\epsilon_2 \sqrt{k^2 + \lambda^2} - p_k k)(\epsilon_1 \sqrt{k^2 + \lambda^2} - p_k k)} \right\},$$

$$B_{\omega_{\min}} = - \int_{k=k_{\min}}^{\Delta\epsilon} \frac{k^2 d|k| d\sigma_k}{|k|} \left\{ \frac{m^2}{(\epsilon_2 |k| - p_k k)^2} + \frac{m^2}{(\epsilon_1 |k| - p_k k)^2} + \frac{2p_1 p_2}{(\epsilon_2 |k| - p_k k)(\epsilon_1 |k| - p_k k)} \right\}. \quad (32.45)$$

where $d\Omega_k$ is the element of solid angle about the photon momentum $\mathbf{k}$
 Let us now make use of the identity [see (42.11)]

$$\frac{1}{(e_2 \sqrt{k^2 + \lambda^2} - p_2 k)(e_1 \sqrt{k^2 + \lambda^2} - p_1 k)} = \frac{1}{2} \int_{-1}^1 \frac{dz}{(e_2 \sqrt{k^2 + \lambda^2} - p_2 k)^2}, \quad (32.46)$$

where

$$p_z = \frac{1}{2}(1+z)p_1 + \frac{1}{2}(1-z)p_2, \quad e_z = \frac{1}{2}(1+z)e_1 + \frac{1}{2}(1-z)e_2.$$

Then

$$B_\lambda = - \int_{|\mathbf{k}|=0}^{|\mathbf{k}|=\infty} \frac{k^2 d|\mathbf{k}| d\Omega_k}{\sqrt{k^2 + \lambda^2}} \left\{ \frac{m^2}{(e_1 \sqrt{k^2 + \lambda^2} - p_1 k)^2} + \frac{m^2}{(e_2 \sqrt{k^2 + \lambda^2} - p_2 k)^2} + \right. \\ \left. + p_1 p_2 \int_{-1}^1 \frac{dz}{(e_z \sqrt{k^2 + \lambda^2} - p_z k)^2} \right\}.$$

Since

$$\int \frac{d\Omega_k}{(e \sqrt{k^2 + \lambda^2} - p k)^2} = \frac{4\pi}{k^2 (e^2 - p^2) + \lambda^2 e^2},$$

we have

$$B_\lambda = -4\pi \int_0^\infty \frac{k^2 d|\mathbf{k}|}{\sqrt{k^2 + \lambda^2}} \left\{ \frac{m^2}{k^2 (e_1^2 - p_1^2) + \lambda^2 e_1^2} + \frac{m^2}{k^2 (e_2^2 - p_2^2) + \lambda^2 e_2^2} + \right. \\ \left. + p_1 p_2 \int_{-1}^1 \frac{dz}{k^2 (e_z^2 - p_z^2) + \lambda^2 e_z^2} \right\}. \quad (32.47)$$

It is easy to see that

$$\int_0^\infty \frac{k^2 d|\mathbf{k}|}{[k^2 (e^2 - p^2) + \lambda^2 e^2] \sqrt{k^2 + \lambda^2}} = \frac{1}{e^2 - p^2} \ln \frac{2\Delta e}{\lambda} - \frac{1}{2} \frac{1}{e^2 - p^2} \frac{e}{p} \ln \frac{e + |p|}{e - |p|}. \quad (32.48)$$

Therefore

$$B_\lambda = 4\pi \left\{ - \left[2 + (p_1 p_2 - e_1 e_2) \int_{-1}^1 \frac{dz}{e_z^2 - p_z^2} \right] \ln \frac{2\Delta e}{\lambda} + \frac{1}{2} \left[\frac{e_1}{|p_1|} \ln \frac{e_1 + |p_1|}{e_1 - |p_1|} + \right. \right. \\ \left. \left. + \frac{e_2}{|p_2|} \ln \frac{e_2 + |p_2|}{e_2 - |p_2|} + (p_1 p_2 - e_1 e_2) \int_{-1}^1 \frac{dz}{e_z^2 - p_z^2} \frac{e_z}{p_z} \ln \frac{e_z + |p_z|}{e_z - |p_z|} \right] \right\}. \quad (32.49)$$

Introducing in place of z , the new variable of integration ξ given by

$$\xi = \frac{|p_z|}{v e_z} = \sqrt{\cos^2 \frac{\theta}{2} + z^2 \sin^2 \frac{\theta}{2}},$$

we can rewrite B_λ in the form

$$B_\lambda = 4\pi \left\{ 2(2\Phi \coth 2\Phi - 1) \ln \frac{2\Delta e}{\lambda} + \frac{1}{v} \ln \frac{1+v}{1-v} - \right. \\ \left. - \frac{1-v^2}{v \sin \frac{\theta}{2}} \coth 2\Phi \int_{\cosh \frac{\theta}{2}}^1 \ln \frac{1+v\xi}{1-v\xi} \frac{d\xi}{(1-v^2\xi^2) \sqrt{\xi^2 - \cos^2 \frac{\theta}{2}}} \right\}, \quad (32.50)$$

where Φ is related to θ by the expression

$$q^2 = 4m^2 \sinh^2 \Phi. \quad (32.51)$$

Since $q_4 = 0$, we have $q^2 = q^2 = 4p^2 \sin^2 \frac{\theta}{2}$, and

$$\sinh \Phi = \frac{v}{\sqrt{1-v^2}} \sin \frac{\theta}{2}. \quad (32.51')$$

Setting $\lambda = 0$ in (32.47) and integrating over $|\mathbf{k}|$ from $|\mathbf{k}| = \omega_{\min}$ to $|\mathbf{k}| = \Delta e$, we obtain

$$B_{\min} = -4\pi \left\{ 2 + (p_1 p_2 - e_1 e_2) \int_{-1}^1 \frac{dz}{e_z^2 - p_z^2} \right\} \int_{\min}^\infty \frac{d|\mathbf{k}|}{|\mathbf{k}|}.$$

Since

$$\int_{-1}^1 \frac{dz}{e_z^2 - p_z^2} = \frac{4\Phi}{m^2 \sinh 2\Phi}$$

and

$$p_1 p_2 = e_1 e_2 = -m^2 \cosh 2\Phi,$$

we finally obtain

$$B_{n, \text{min}} = 8\pi (2\Phi \cosh 2\Phi - 1) \ln \frac{\Delta \epsilon}{\omega_{\text{min}}}. \quad (32.52)$$

Comparing (32.50) and (32.52) we find the desired relation between λ and ω_{min} , namely

$$2(1 - 2\Phi \cosh 2\Phi) \ln \frac{2\omega_{\text{min}}}{\lambda} = \frac{1}{v} \ln \frac{1+v}{1-v} - \frac{(1-v^2)}{v \sin \frac{\theta}{2}} \cosh 2\Phi \int_0^1 \ln \frac{1+v^2}{1-v^2} \frac{d^2 \epsilon}{(1-v^2)^2 \sqrt{\epsilon^2 - \cos^2 \frac{\theta}{2}}}. \quad (32.53)$$

In the limiting case of low electron energy ($v \ll 1$) (32.53) reduces to

$$\ln 2\omega_{\text{min}} = \ln \lambda + \frac{5}{6}. \quad (32.54)$$

Inserting (32.50) into (32.42), we obtain the following expression for the differential cross section for scattering of an electron by a Coulomb field with the emission of a photon whose energy is $\Delta \epsilon$ or less (where $\Delta \epsilon \ll \epsilon$):

$$d\sigma' = \left(\frac{Z^2 \alpha}{2m v^2 \sin^2 \frac{\theta}{2}} \right)^2 (1 - v^2) \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \frac{\alpha}{\pi} \left\{ 2(2\Phi \cosh 2\Phi - 1) \ln \frac{2\Delta \epsilon}{\lambda} + \frac{1}{v} \ln \frac{1+v}{1-v} - \frac{1-v^2}{v \sin \frac{\theta}{2}} \cosh 2\Phi G(v, \theta) \right\} d\theta, \quad (32.55)$$

where

$$G(v, \theta) = \int_0^1 \ln \frac{1+v^2}{1-v^2} \frac{d^2 \epsilon}{(1-v^2)^2 \sqrt{\epsilon^2 - \cos^2 \frac{\theta}{2}}}.$$

§ 33. Disintegration of an Electron-Positron Pair into Photons. Pair Production by Photons.

1. Two-Photon Annihilation

When an electron collides with a positron, they may disintegrate into photons. If both particles are free, disintegration into one photon is forbidden by conservation of energy and momentum. Two-photon annihilation is

described by the second-order matrix $\gamma_S^{(2)}$, which we have already studied.

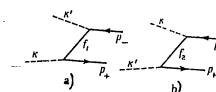


Fig. 40

Let us construct the diagrams for the matrix element corresponding to this process. These are shown in Fig. 40. The four-momenta of the electron and positron are denoted by p_- and p_+ , respectively, the four-momenta of the photons are denoted by k and k' , and f_1 and f_2 are the four-momenta of the virtual electrons. As previously, the sum of the momenta of the segments to the right and to the left of each vertex are equal, so that

$$\begin{aligned} f_1 &= p_- - k' = k - p_+, \\ f_2 &= p_- - k = k' - p_+. \end{aligned} \quad (33.1)$$

It is simple to obtain the expression for the matrix element from the general rules [see (24.7)] this is

$$S_{f_1 f_2}^{(2)} = \frac{-ie^2}{2\sqrt{\omega\omega'}} \bar{u} \left\{ \frac{f_1 \not{p}_+ - m}{m^2 \epsilon_1} \not{\epsilon}' + \not{\epsilon} \frac{f_2 \not{p}_- - m}{m^2 \epsilon_2} \right\} u (2\pi)^4 \delta(p_+ + p_- - k - k'). \quad (33.2)$$

Here $\bar{u}$ is the amplitude of the electron wave function, and u is the amplitude of the negative-frequency wave function corresponding to the positron state.

We note that the structure of (33.2) is exactly the same as that of the S matrix element (29.5) for scattering a photon by an electron: if we perform the substitutions

$$\begin{aligned} p_1 &= p_-; & k_1 &= -k'; & e_1 &= e', \\ p_2 &= -p_+; & k_2 &= k; & e_2 &= e, \end{aligned} \quad (33.3)$$

in the latter, we obtain (33.2)

Let us now find the cross section for two-photon annihilation. The general expression for the cross section is

$$d\sigma = \frac{e^4}{4\omega\omega'} |\bar{u} Q u|^2 \frac{d^3 k d^3 k'}{(2\pi)^2} \delta(p_+ + p_- - k - k') \delta(\epsilon_+ + \epsilon_- - \omega - \omega')$$

(Q is the operator in the braces in Equation (33.2). When the electron is at rest, the current density is

$$J = v_+$$

and

$$\left| \frac{\partial(\omega + \omega')}{\partial \omega} \right| = \frac{(\epsilon_+ + m)\omega - p_+ k}{\omega\omega'}.$$

Averaging over electron spins and summing over photon polarizations, we obtain

$$d\sigma = \frac{1}{16} r_0^2 \frac{\omega^2 d\Omega}{[m(\epsilon + \epsilon_0) - \mathbf{p} \cdot \mathbf{k}] v_+} \text{Sp } F,$$

where F is the same operator as in (29.20). Equations (29.21) with (33.3) can be used for calculating $\text{Sp } F$.

We present the expression for the integral cross section for two-photon annihilation when the electron is at rest¹⁾:

$$\sigma = \pi r_0^2 \frac{1}{x+1} \left[\frac{x^2+4x+1}{x^2-1} \ln(x + \sqrt{x^2-1}) - \frac{x+3}{\sqrt{x^2-1}} \right], \quad (33.4)$$

where

$$x = \frac{\epsilon_+}{m}.$$

It is seen from (33.4) that the cross section increases as the energy increases. At very high positron energies ($\epsilon_+ \gg m$)

$$\sigma = \pi r_0^2 \frac{m}{\epsilon_+} \left(\ln \frac{2\epsilon_+}{m} - 1 \right). \quad (33.5)$$

At low positron energies the cross section is inversely proportional to the positron velocity,

$$\sigma = \frac{\pi r_0^2}{v_+}. \quad (33.6)$$

2. Positronium Decay

These results can be applied to the problem of annihilation of positronium, i.e., the atomic system consisting of an electron and a positron. The ground state (S -state) of positronium can occur in two forms, since it involves two spin-1/2 particles: 1) that with angular momentum 0 (parapositronium) and 2) that with angular momentum 1 (orthopositronium). Since there exists no two-photon system with angular momentum 1 (see Section 6), orthopositronium cannot decay into two photons²⁾. In the second-order approximation, its decay probability is zero. Of the four possible orientations of the electron and positron spins, three belong to orthopositronium, and one to parapositronium; therefore the decay probability w_{para} of parapositronium is four times the mean probability w ³⁾:

$$w = \frac{1}{4} w_{\text{para}} + \frac{3}{4} w_{\text{ortho}} = \frac{1}{4} w_{\text{para}}$$

¹⁾ P. Dirac, Proc. Cambridge Phil. Soc. 26, 361 (1930); I. Tamm, Z. Physik 62, 545 (1930).

²⁾ Furthermore, since the system has odd charge parity, orthopositronium cannot decay into any even number of photons.

³⁾ I. Pomeranchuk, Proc. Acad. Sci. USSR 60, 218 (1947).

The magnitude of w is simply related to the cross section for two-photon pair annihilation for slow positrons as given by (33.6). Since the momenta of the particles in positronium are small compared with the momenta of the decay radiation, the positronium wave function

$$\psi = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0},$$

(where $a_0 = 2a_B$ is the radius of the positronium atom, a_B is the radius of the hydrogen atom, and r is the distance between the particles) may be considered constant

$$\psi \approx \psi(0) = \frac{1}{\sqrt{\pi a_0^3}} \quad \left(\frac{1}{a_0} \ll m \right).$$

Therefore the electron and positron may be considered free particles with zero momenta, but with different normalizing constants (normalized not for unit volume, but for a volume πa_0^3).

Thus the probability w is related to the slow-positron annihilation cross section by

$$w = J\sigma = |\psi(0)|^2 (v_+ a_0)^2. \quad (33.7)$$

Using (33.6), we obtain

$$w_{\text{para}} = 4w = \frac{1}{2} \frac{r_0^2}{a_0^2} = \frac{1}{2} \alpha^6 m = 0.8 \cdot 10^{10} \text{ sec}^{-1} \quad (33.8)$$

3. Three-Photon Orthopositronium Decay

The problem of orthopositronium decay necessitates the consideration of three-photon electron-positron annihilation. We note that parapositronium (in its ground state) cannot decay into three photons.

Indeed, as was shown in Section 19, paragraph 3, the 1S positronium state has even charge parity and can therefore decay only into an even number of photons. (The 3S state, which is the orthopositronium ground state, on the contrary, has odd charge parity.)

Since three-photon parapositronium decay has zero probability, the orthopositronium decay probability is

$$w_{\text{ortho}} = \frac{4}{3} w,$$

where w is the probability averaged over spin orientations. The determination of w necessitates only the consideration of free-particle annihilation with $v_+ = 0$ into three photons, after which we may use (33.7).

Figure 41 shows one of the diagrams illustrating three-photon pair annihilation. The remaining diagrams differ from the ones shown only by permutations of the photon momenta. The matrix element of the third-order matrix $S^{(3)}$ which corresponds to this process can be written

$$S_{1 \rightarrow f}^{(0)} = \frac{e^2}{\sqrt{8\pi\omega_0\omega_1\omega_2}} \left\{ \frac{1}{\omega_1} \frac{1}{r^2 + m^2} \hat{e}_1 \cdot \frac{1}{r^2 + m^2} \hat{e}_2 \cdot \frac{1}{r^2 + m^2} \hat{e}_3 \cdot u + \dots \right\} \times \\ \times (2\pi)^4 \delta(p_- + p_+ - k_1 - k_2 - k_3) \quad (33.9)$$

(the three dots correspond to similar terms in which the vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$ and k_1, k_2, k_3 have been permuted).

The differential probability for three-photon orthopositronium decay as given by (33.9) averaged over electron and positron spins and summed over photon polarizations is¹⁾

$$dw_{\text{ortho}} = \frac{2}{3} \frac{\alpha^3}{(2\pi)^3} \frac{1}{a_0^3} \{ (1 - n_1 n_2)^2 + (1 - n_2 n_3)^2 + (1 - n_3 n_1)^2 \} \times \\ \times dk_1 dk_2 dk_3 \delta(k_1 + k_2 + k_3) \delta(\omega_1 + \omega_2 + \omega_3 - 2m), \quad (33.10)$$

where $\hat{n}_1, \hat{n}_2$ and $\hat{n}_3$ are unit vectors in the directions of motion of the photons.

The total probability is

$$w_{\text{ortho}} = \frac{1}{6} \int dw_{\text{ortho}}$$

where the integration is taken over all photon momentum space, and the factor 1/6 takes account of the indistinguishability of states which differ by permutations of the photon momenta. Integrating this expression leads to

$$w_{\text{ortho}} = \frac{2}{9\pi} (\pi^2 - 9) \alpha \frac{r_0^2}{a_0^3} \approx \frac{1}{8.137} w_{\text{para}} = 0.71 \cdot 10^7 \text{ sec}^{-1}. \quad (33.11)$$

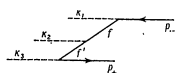


Figure 41

Integration of (33.10) over all variables other than ω gives the photon spectrum¹⁾

$$F(\omega) = \frac{4}{9\pi} \left(\frac{r_0}{a_0} \right)^3 \left\{ \frac{\omega(m-\omega)}{(2m-\omega)^2} \ln \frac{m-\omega}{m} + \frac{2m-\omega}{\omega} \ln \frac{m-\omega}{m} + \right. \\ \left. + \frac{2(m-\omega)m}{\omega^2} \ln \frac{m-\omega}{m} \right\} \quad (33.12)$$

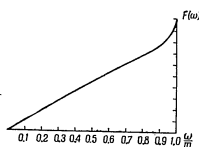


Figure 42

¹⁾ A. Ore and J. Powell, Phys. Rev. 75, 1696 (1949).

$$\left(\int_0^m F(\omega) d\omega = w_{\text{ortho}} \right).$$

The maximum photon energy allowed by the conservation laws is m . Figure 42 is a graph of (33.12).

4. Single-Photon Pair Annihilation.

When a positron collides with a bound electron (for instance a K-electron) single-photon annihilation may take place. Such a process is given by the first order matrix element

$$S_{1 \rightarrow f}^{(1)} = \frac{-e}{\sqrt{2\omega}} \int \bar{\psi}_2 \hat{e} \psi_1 d^4x, \quad (33.13)$$

where ψ_1 is the electron wave function, and ψ_2 is the negative-frequency wave function corresponding to the positron.

This matrix element is calculated in the same way as that for the photoelectric effect. We present the results obtained in the free-positron approximation [ψ_2 in (33.13) is a plane wave]. The annihilation cross section for the two K-electrons is¹⁾

$$\sigma = 4\pi r_0^2 Z^2 \alpha^4 \frac{1}{(x+1)^2 \sqrt{x^2-1}} \left[x^2 + \frac{2}{3}x + \frac{4}{3} - \frac{x+2}{\sqrt{x^2-1}} \ln(x + \sqrt{x^2-1}) \right],$$

where

$$x = \frac{e_+}{m}. \quad (33.14)$$

For low positron energies ($|p_+| \ll m$)

$$\sigma = \frac{4\pi}{3} r_0^2 Z^2 \alpha^4 v_+. \quad (33.15)$$

Thus, the cross section for single-photon annihilation, as opposed to that for two-photon annihilation, decreases as the positron energy decreases. At low energies the ratio of the cross section for single-photon annihilation to that for two-photon annihilation is extremely small. This ratio is a maximum when $e_+ \sim 10m$ (for $Z = 82$ it is about 20 per cent), but at such energies the absolute values of the cross sections are small.

5. Pair Production by a Photon in an External Field

Let us now go on to a consideration of electron-positron pair production. Pair production by two photons

¹⁾ See W. Heitler, The Quantum Theory of Radiation (State Tech. Press, 1940) p. 229.

is the inverse of the above two-photon pair annihilation process, and is not of great interest because of the low photon densities usually encountered. Quite important, however, is pair creation by a photon colliding with a charged particle. If this charged particle is heavy (say a nucleus), then the problem is very similar to that of bremsstrahlung in the field of the nucleus.

The diagrams corresponding to this process are shown in Fig. 43. The matrix element can be written

$$S_{fi}^{(0)} = \frac{-ie^2}{\sqrt{2\omega}} \bar{u} \left\{ \hat{a}_\mu \frac{f_1 - m}{f_1^2 + m^2} \hat{e} + \hat{e} \frac{f_2 - m}{f_2^2 + m^2} \hat{a}_\mu \right\} v, \quad (33.16)$$

where

$$f_1 = -p_+ + k, \quad f_2 = p_- - k.$$

We note that the matrix element (31.4) for bremsstrahlung is transformed into (33.16) by the substitutions

$$p_1 = -p_+, \quad u_2 = u, \quad k \rightarrow -k, \\ p_2 = p_-, \quad u_1 = v, \quad f_1 \leftrightarrow f_2$$

and

$$q = p_+ + p_- - k.$$

Therefore, a calculation of the cross section for pair production differs from that for the bremsstrahlung cross section only in the different density of final states, which is now given by

$$\frac{d^3 p_+ d^3 p_-}{(2\pi)^6},$$

and the different value for the current density, which is now unity.

In analogy with (31.19), we obtain the following expression for the differential cross section for pair production by a photon in a Coulomb field¹⁾:

$$d\sigma = \frac{Z^2 e^2}{(2\pi)^2} \frac{1}{\omega^3} \frac{|p_+||p_-|}{\omega^3} \frac{d\epsilon_+ d\epsilon_-}{q^4} \frac{d\omega_-}{m^2} \left\{ 4 \left[k_1 \frac{\epsilon_+ p_-}{x_1} + \frac{\epsilon_- p_+}{x_2} \right]^2 - \right. \\ \left. - q^2 \left[k_1 \frac{p_-}{x_1} - \frac{p_+}{x_2} \right]^2 - \frac{2\omega^2}{x_1 x_2} (k_1 p_- + p_+)^2 \right\}, \quad (33.17) \\ m^2 x_1 = -2k p_+, \quad m^2 x_2 = -2k p_-.$$

¹⁾ H. Bethe and W. Heitler, Proc. Roy. Soc. 146, 83 (1934).

Integration of (33.17) over angles gives the cross section for pair production with a given positron energy ϵ_+ (we note that $\epsilon_- = \omega - \epsilon_+$). Let us denote this by $d\sigma_\epsilon$:

$$d\sigma_\epsilon = \frac{4\pi}{\omega^3} \frac{|p_+||p_-|}{\omega^3} d\epsilon_+ \left\{ -\frac{4}{3} - 2\epsilon_+ \epsilon_- \frac{p_+^2 + p_-^2}{p_+^2 p_-^2} + \right. \\ \left. + m^2 \left(\frac{\epsilon_+ \epsilon_-}{|p_+||p_-|} + \frac{\epsilon_- \epsilon_+}{|p_+||p_-|} - \frac{\epsilon_+ \epsilon_-}{|p_+||p_-|} \right) + L \left[\frac{\omega^2}{|p_+||p_-|} (\epsilon_+^2 \epsilon_-^2 + p_+^2 p_-^2) - \right. \right. \\ \left. \left. - \frac{8}{3} \frac{\epsilon_+ \epsilon_-}{|p_+||p_-|} - \frac{m^2 \omega}{2|p_+||p_-|} \left(\frac{\epsilon_+ \epsilon_- - p_-^2}{|p_-|^3} \epsilon_- + \frac{\epsilon_+ \epsilon_- - p_+^2}{|p_+|^3} \epsilon_+ + \frac{2\omega \epsilon_+ \epsilon_-}{|p_+||p_-|^3} \right) \right] \right\}, \quad (33.18) \\ \zeta_\pm = 2 \ln \frac{\epsilon_\pm + |p_\pm|}{m}; \quad L = 2 \ln \frac{\epsilon_+ \epsilon_- + |p_+||p_-| + m^2}{m\omega}; \quad \bar{Q} = Z^2 \alpha r_0^2.$$

For high energies (ω, ϵ_+ and $\epsilon_- \gg m$) Equation (33.18) can be written

$$d\sigma_\epsilon = 4\bar{Q} d\epsilon_+ \frac{\epsilon_+^2 + \epsilon_-^2 + \frac{2}{3} \epsilon_+ \epsilon_-}{\omega^3} \left(\ln \frac{2\epsilon_+ \epsilon_-}{\omega m} - \frac{1}{2} \right). \quad (33.19)$$

These equations, as those in the bremsstrahlung problem, are restricted by the condition

$$\xi_+ = \frac{Z\alpha}{v_+} \ll 1; \quad \xi_- = \frac{Z\alpha}{v_-} \ll 1.$$

When $v_+ \ll 1$ and $v_- \ll 1$, the problem can be solved by using the exact nonrelativistic wave functions. In this case (33.18) should be multiplied by

$$f = \frac{(2\pi)^2 \xi_+ \xi_-}{(e^{\pi\xi_+} - 1)(1 - e^{-\pi\xi_-})},$$

where

$$\xi_\pm = \frac{Z\alpha}{v_\pm}.$$

At high energies we must take account of the screening of the nuclear charge by the electron shells. Equation (33.18) is valid so long as

$$\frac{2\epsilon_+ \epsilon_-}{\omega m} \ll 137 Z^{-1/2},$$

as can be seen from considerations similar to those in Section 31. In the other limiting case of complete screening, that is when

$$\frac{2\epsilon_+ \epsilon_-}{\omega m} \gg 137 Z^{-1/2},$$

we obtain an expression similar to (31.23), namely

$$d\sigma = 4\pi d\epsilon_+ \left\{ \frac{1}{\omega^2} \left[\epsilon_+^2 + \epsilon_-^2 + \frac{2}{3} \epsilon_+ \epsilon_- \right] \ln(183 Z^{-1/2}) - \frac{1}{9} \frac{\epsilon_+ \epsilon_-}{\omega^2} \right\}. \quad (33.20)$$

Finally, we present the formula for the total pair creation cross section for high energies. Integration over positron energies gives

$$\sigma = \pi \left(\frac{28}{9} \ln \frac{2\omega}{m} - \frac{218}{27} \right) \quad (33.21)$$

in the absence of screening, and

$$\sigma = \pi \left(\frac{28}{9} \ln(183 Z^{-1/2}) - \frac{2}{27} \right). \quad (33.22)$$

in the case of complete screening.

Figure 44 shows the dependence of the cross section on the positron energy. The abscissa gives the ratio of the kinetic energy $\epsilon_+ - m$ of the positron to the total kinetic energy $\omega - 2m$ of the pair; the ordinate gives $\frac{1}{\pi} \frac{d\sigma}{d\epsilon_+} (\omega - 2m)$. Figure 45 shows the dependence of the total cross section for pair creation on photon energies for various Z . The top curve gives the cross section without screening.

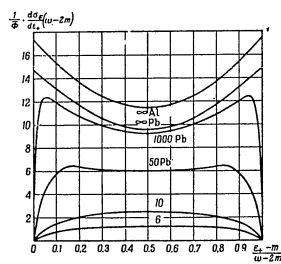


Fig. 44

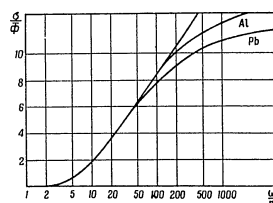


Fig. 45

6. Pair Production in Photon-Electron Collisions

The problem of pair production by the collision of a photon with an electron is similar to that of photon emission in the collision of two electrons (see Section 31). Figure 46 shows two of the eight diagrams which illustrate this process. The remaining diagrams are obtained from these by interchanging the photons and virtual photons and the momenta p_2 and p_1 of the electrons in the final states (p_1 is the initial electron momentum, and p_2 is the positron momentum). The matrix element $S_{if}^{(2)}$ for this process is the same as (31.26) with p_2 replaced by $-p_2$, k by $-k$, and ω_i by the negative-frequency amplitude ω . We shall present the integral cross section for pair creation by a photon and an electron at rest in two limiting cases.

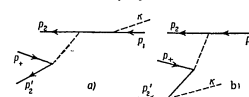


Fig. 46

Near the pair-production threshold the cross-section

$$\sigma = \alpha r_0^2 \pi \frac{\sqrt{s}}{4 \cdot 3^3} \left(\frac{\omega}{m} - 4 \right)^2. \quad (33.23)$$

(We note that $\omega = 4m$ is the lowest photon energy allowed by the conservation laws.)

In the ultrarelativistic case ($\omega \gg m$)¹⁾

$$\sigma = \alpha r_0^2 \left(\frac{28}{9} \ln \frac{2\omega}{m} - 10.8 \right). \quad (33.24)$$

Equation (33.24) differs from the cross section for pair production by a photon in the field of a nucleus with $Z = 0$ [see (33.21)] only by the coefficient of the logarithm.

In all the above calculations for pair production we did not take account of the interaction between the particles produced (electrons and positrons). A more exact solution of this problem with this interaction taken into account is relatively complicated, and necessitates calculation of higher approximations in the S matrix. The interaction is most important, however, at low relative velocities of the produced particles. In this case the difference reduces to an increase (due to the attraction between the electron and positron) in the probability for finding the particles at a single point. If we write the differential cross section for pair production without the interaction in the form

$$d\sigma = A(P, p) dP dp, \quad (33.25)$$

where P is the momentum of the center of mass of the produced pair and p is the relative momentum of the particles, inclusion of the interaction of small p reduces²⁾ to multiplication of (33.25) by $\psi^2(0)$, so that

$$d\sigma = \psi^2(0) A(P, p) dP dp, \quad (33.26)$$

where $\psi(r)$ is the wave function of the relative motion (r is the relative radius vector) normalized for a unit momentum interval

$$\psi^2(0) = \frac{2\pi\alpha}{v(1 - e^{-2\pi\alpha/v})}. \quad (33.27)$$

¹⁾ E. Nemirowsky, J. Phys. 11, 94 (1947).

²⁾ Votruba, Phys. Rev. 73, 1408 (1948).

³⁾ A. Sakharov, J. Exptl.-Theoret. Phys. 18, 631 (1948).

(v is the relative velocity).

§ 34. Photon Scattering by a Bound Electron. Emission of Two Photons.

1. Dispersion Equation.

We have so far applied the matrix $\gamma S^{(2)}$ to the problem of photon scattering by an electron only for free-electron states. For bound electrons, when the wave functions of the stationary states are not plane waves, the matrix $\gamma S^{(2)}$ still has the form of (28.2), but $\bar{g}^E$ is no longer given by (18.31) and (18.33). We must now use the expressions

$$\frac{1}{2} S^{(2)}(2, 1) = \begin{cases} \sum_n \bar{\psi}_n^{(+)}(1) \psi_n^{(+)}(2) & (t_2 > t_1), \\ -\sum_n \bar{\psi}_n^{(-)}(1) \psi_n^{(-)}(2) & (t_2 < t_1) \end{cases} \quad (34.1)$$

[see (18.28') and (18.30); $\psi_n^{(+)}$ and $\psi_n^{(-)}$ are positive and negative-frequency wave functions, respectively] which in general cannot be given in closed form.

If we separate out the time dependence from ψ_n , writing

$$\psi_n(1) = \psi_n(r_1) e^{-i\epsilon_n t_1},$$

then $\bar{g}^E$ can also be written

$$\frac{1}{2} S^{(2)}(2, 1) = \frac{1}{2\pi i} \sum_n \bar{\psi}_n(r_1) \psi_n(r_2) \int_{-\infty}^{\infty} \frac{e^{i\omega(t_2-t_1)}}{\epsilon_n + \omega} d\omega, \quad (34.1')$$

where the pole of the integrand is treated according to the general rule formulated in Section 18. From (29.2) and (34.1) we obtain the following expression for $\gamma S^{(2)}$:

$$\begin{aligned} \gamma S^{(2)} = e^2 \int_{-\infty}^{\infty} dt_2 \int_{-\infty}^{t_2} dt_1 \sum_n \bar{\psi}(2) \hat{A}(2) \psi_n^{(+)}(2) dr_2 \times \\ \times \int \bar{\psi}_n^{(+)}(1) \hat{A}(1) \psi(1) dr_1 - e^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \sum_n \bar{\psi}(2) \hat{A}(2) \psi_n^{(-)}(2) dr_2 \times \\ \times \int \bar{\psi}_n^{(-)}(1) \hat{A}(1) \psi(1) dr_1. \end{aligned} \quad (34.2)$$

Let us consider photon scattering by an electron. We shall write the stationary-state electron wave functions in the form

$$\psi(x) = \psi(r) e^{-iEt},$$

and the photon potentials in the form

$$A(x) = A(r) e^{-i\omega t} = \frac{e}{\sqrt{2\omega}} e^{i(kr - \omega t)}.$$

The initial states of the electron and photon shall be denoted by the indices 1, and the final states by the indices 2, as previously. The frequency of the scattered photon may be the same as that of the incident one, or it may differ from it (combined scattering). Inserting (34.1) into the general formula (29.3) for the matrix element, we obtain

$$\begin{aligned} S_{1 \rightarrow 2}^{(2)} = e^2 \sum_n \left\{ (a_2^* \gamma_{2n}^{(+)})(a_1 \gamma_{n1}^{(+)} \int_{-\infty}^{\infty} e^{i(\epsilon_2 - \omega_2 - \epsilon_n^{(+)} t_2) t_2} dt_2 \int_{-\infty}^{t_2} e^{i(\epsilon_1^{(+)} - \omega_1 - \epsilon_n^{(+)} t_1) t_1} dt_1 + \right. \\ \left. + (a_1 \gamma_{2n}^{(+)})(a_2^* \gamma_{n1}^{(+)} \int_{-\infty}^{\infty} e^{i(\epsilon_2 - \omega_2 - \epsilon_n^{(+)} t_2) t_2} dt_2 \int_{-\infty}^{t_2} e^{i(\epsilon_1^{(+)} - \omega_1 - \epsilon_n^{(+)} t_1) t_1} dt_1 - \right. \\ \left. - (a_2^* \gamma_{2n}^{(-)})(a_1 \gamma_{n1}^{(-)} \int_{-\infty}^{\infty} e^{i(\epsilon_2 - \omega_2 - \epsilon_n^{(-)} t_2) t_2} dt_2 \int_{-\infty}^{t_2} e^{i(\epsilon_1^{(-)} - \omega_1 - \epsilon_n^{(-)} t_1) t_1} dt_1 - \right. \\ \left. - (a_1 \gamma_{2n}^{(-)})(a_2^* \gamma_{n1}^{(-)} \int_{-\infty}^{\infty} e^{i(\epsilon_2 - \omega_2 - \epsilon_n^{(-)} t_2) t_2} dt_2 \int_{-\infty}^{t_2} e^{i(\epsilon_1^{(-)} - \omega_1 - \epsilon_n^{(-)} t_1) t_1} dt_1 \right\}, \end{aligned} \quad (34.3)$$

where

$$\begin{aligned} (a_1)_{n1}^{(+)} &= \int \bar{\psi}_2(r) \hat{A}_1(r) \psi_n^{(+)}(r) dr = \frac{-ie_1}{\sqrt{2\omega_1}} \int \bar{\psi}_2 e^{i(k_1 r - \omega_1 t_1)} \psi_n^{(+)} dr, \\ (a_2)_{2n}^{(+)} &= \int \bar{\psi}_2 \hat{A}_2 \psi_n^{(+)} dr = \frac{-ie_2}{\sqrt{2\omega_2}} \int \bar{\psi}_2 e^{-i(k_2 r - \omega_2 t_2)} \psi_n^{(+)} dr, \\ (a_1)_{n1}^{(-)} &= \int \bar{\psi}_2 \hat{A}_1 \psi_n^{(-)} dr = \frac{-ie_1}{\sqrt{2\omega_1}} \int \bar{\psi}_2 e^{i(k_1 r - \omega_1 t_1)} \psi_n^{(-)} dr, \\ (a_2)_{2n}^{(-)} &= \int \bar{\psi}_2 \hat{A}_2 \psi_n^{(-)} dr = \frac{-ie_2}{\sqrt{2\omega_2}} \int \bar{\psi}_2 e^{-i(k_2 r - \omega_2 t_2)} \psi_n^{(-)} dr. \end{aligned} \quad (34.4)$$

The double integration over time can be performed in (34.3). For the first integration, the value of the integral at the lower limit may be considered zero (the undefined quantity $e^{i\omega\infty}$ can be eliminated by multiplying the integrand by $\bar{g}^{\lambda E}$ and then letting λ approach zero). Then the second integration leads to a δ -function and therefore to energy conservation. Thus

$$\begin{aligned} S_{1 \rightarrow 2}^{(2)} = ie^2 \sum_n \left\{ \frac{(a_2^* \gamma_{2n}^{(+)})(a_1 \gamma_{n1}^{(+)} + (a_1 \gamma_{2n}^{(+)})(a_2^* \gamma_{n1}^{(+)} + (a_2^* \gamma_{2n}^{(-)})(a_1 \gamma_{n1}^{(-)} + \right. \\ \left. + (a_1 \gamma_{2n}^{(-)})(a_2^* \gamma_{n1}^{(-)}) \right\} 2\pi \delta(\epsilon_1 + \omega_1 - \epsilon_2 - \omega_2), \end{aligned} \quad (34.5)$$

We shall also write $\bar{S}_{1 \rightarrow 2}^{(2)}$ in the form of (28.17), namely

$$S_{1 \rightarrow 2}^{(2)} = -2\pi i \delta(\epsilon_1 + \omega_1 - \epsilon_2 - \omega_2) U_{1 \rightarrow 2}.$$

Going on to a consideration of this case, we shall assume that $|\epsilon_1 - m|$, $|\epsilon_2 - m|$, ω_1 and ω_2 are small compared to m . In addition, we shall assume that only those states whose energy is nonrelativistic, i.e., for which $|\epsilon_n - m| \ll m$, contribute significantly to (34.5). Then a major simplification is possible for that part of the sum which refers to negative frequencies. First, we shall replace the denominator in each term by $2m$. Denoting this part of the sum by Σ_{-} , we have

$$\Sigma_{-} = \frac{1}{2m} \sum_n [(a_2^{\dagger})_{2n}^{\dagger} (a_1)_{1n}^{\dagger} + (a_1)_{2n}^{\dagger} (a_2^{\dagger})_{1n}^{\dagger}].$$

This can be transformed to a sum over all states (including positive frequencies). To do this, we replace $\psi_n^{(-)}$ in the expression for $(a_2^{\dagger})_{2n}^{\dagger}$ by

$$\frac{m-H}{2m} \psi_n^{(-)},$$

where H is the electron Hamiltonian. In our approximation

$$H = \beta m, \quad H \psi_n^{(+)} = -\frac{1}{2m} \nabla^2 \psi_n^{(+)},$$

and therefore

$$\frac{m-H}{2m} \psi_n^{(-)} = \psi_n^{(-)}, \quad (m-H) \psi_n^{(+)} = 0$$

and

$$\Sigma_{-} = \frac{1}{2m} \sum_n \left\{ \int \psi_n^{\dagger} A_2^{\dagger} \alpha \frac{(1-\beta)}{2} \psi_n dr \int \psi_n^{\dagger} \alpha A_1 \psi_1 dr + \right. \\ \left. + \int \psi_n^{\dagger} A_1 \alpha \frac{(1-\beta)}{2} \psi_n dr \int \psi_n^{\dagger} \alpha A_2^{\dagger} \psi_1 dr \right\},$$

where the sum is taken over both signs of the frequency. We can now use the completeness of the set of functions ψ_n (that is, the matrix multiplication rule), so that

$$\Sigma_{-} = \frac{1}{2m} \int \psi_2^{\dagger} \left[(A_2^{\dagger} \alpha) \frac{(1-\beta)}{2} (\alpha A_1) + (A_1 \alpha) \frac{(1-\beta)}{2} (\alpha A_2^{\dagger}) \right] \psi_1 dr.$$

Using the commutation relations

$$\beta \alpha = -\alpha \beta, \\ (A_2^{\dagger} \alpha) (\alpha A_1) + (A_1 \alpha) (\alpha A_2^{\dagger}) = 2 A_2^{\dagger} A_1$$

and the fact that in our approximation

$$\psi_1 = \psi_1^{+},$$

we obtain

$$\Sigma_{-} = \frac{1}{m} (A_2^{\dagger} A_1)_{21} = \frac{1}{m} \int \psi_2^{\dagger} A_2^{\dagger} A_1 \psi_1 dr = \frac{e_1 e_2}{2 \sqrt{\omega_1 \omega_2}} \frac{1}{m} \int \psi_2^{\dagger} \psi_1 e^{i(\omega_1 - \omega_2)r} dr. \quad (34.6)$$

In the first part of (34.6) referring to positive frequencies, the energy difference $|\epsilon_1 - \epsilon_n| \ll m$ enters the denominators, and therefore, they cannot be simplified. According to (14.7) in the nonrelativistic approximation αA in the numerators can be replaced by

$$\alpha A = \frac{1}{m} A p + \mu \text{curl} A,$$

where $\mu = \frac{e}{2m} \sigma$ is the spin magnetic moment of the electron. Thus, according to (34.5) and (34.6), we obtain

$$U_{i \rightarrow f} = \frac{e^2}{m} \left(\frac{1}{m} \sum_n \left\{ \frac{(p' A_2)_{2n} (p' A_1)_{n1}}{\epsilon_1 - \epsilon_n + \omega_1} + \frac{(p' A_1)_{2n} (p' A_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right\} + (A_2^{\dagger} A_1)_{21} \right), \quad (34.7)$$

where $p' = p + \frac{e}{2} \sigma \text{curl} A$, and the parentheses indicate the matrix elements of the corresponding operators according to

$$(\quad)_{nq} = \int \psi_n^{\dagger} (\quad) \psi_q dr,$$

where $\psi_n^{\pm}$, $\psi_q^{\pm}$ are the nonrelativistic wave functions (therefore, the index $(+)$ is dropped from the ϵ_n).

The matrix element (34.7) is exactly the same as the result as could have been obtained by starting with the nonrelativistic interaction energy operator of the electron and electromagnetic field [see (14.7)]

$$V = -\frac{ep'}{m} A + \frac{e^2}{2m} A^2.$$

In this operator the term quadratic in A is responsible for "direct transitions"; in the first approximation of perturbation theory it gives the last term in (34.7). The linear term accounts for transitions through "intermediate states", and in the second approximation it gives the sum in (34.7).

If the photon wavelength is large compared with the dimensions of the scatterer, then the A_1 and A_2 can be considered constant ($e^{ikr} \approx \frac{r}{R} e^{ikR}$, where R is the radius vector of the center of the atom, for which we may set $R = 0$). Then (34.7) becomes

$$U_{i \rightarrow f}^{(0)} = \frac{e^2 \delta^3(\mathbf{r}_i - \mathbf{r}_f) R}{2m \sqrt{\omega_1 \omega_2}} \left\{ \frac{1}{m} \sum_n \left[\frac{(\mathbf{p} \mathbf{e}_1)_{2n} (\mathbf{p} \mathbf{e}_1)_{n1}}{\epsilon_1 - \epsilon_n + \omega_1} + \frac{(\mathbf{p} \mathbf{e}_1)_{2n} (\mathbf{p} \mathbf{e}_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right] + e_1 e_2 \delta_{12} \right\}, \quad (34.8)$$

The last term here differs from zero only when the initial and final electron states are identical ($\omega_1 = \omega_2$; coherent scattering).

Equation (34.8) gives the dispersion law, i.e., the dependence of atomic scattering on the frequency of the light. It can be written in a somewhat different form. The matrix elements of the components of the momenta can be written in terms of the matrix elements of the coordinates of the electron (or in terms of its dipole moment) according to the well known relation

$$(\mathbf{p})_{i\alpha} = im(\epsilon_\alpha - \epsilon_i)(\mathbf{r})_{i\alpha}.$$

The last term in (34.8) can be combined with the first two by using the commutation relations between the momenta and coordinates. We may write, namely,

$$e_1 e_2 \delta_{12} = i[(\mathbf{p} \mathbf{e}_1)(\mathbf{r} \mathbf{e}_2) - (\mathbf{r} \mathbf{e}_1)(\mathbf{p} \mathbf{e}_2)]_{21} = i \sum_n [(\mathbf{p} \mathbf{e}_1)_{2n} (\mathbf{r} \mathbf{e}_2)_{n1} - (\mathbf{r} \mathbf{e}_1)_{2n} (\mathbf{p} \mathbf{e}_2)_{n1}].$$

We may now add the vanishing expression

$$0 = \omega_2 [(\mathbf{r} \mathbf{e}_1)(\mathbf{r} \mathbf{e}_2) - (\mathbf{r} \mathbf{e}_2)(\mathbf{r} \mathbf{e}_1)]_{21} = \omega_2 \sum_n [(\mathbf{r} \mathbf{e}_1)_{2n} (\mathbf{r} \mathbf{e}_2)_{n1} - (\mathbf{r} \mathbf{e}_2)_{2n} (\mathbf{r} \mathbf{e}_1)_{n1}],$$

to (34.8), and write $U_{i \rightarrow f}$ in the form

$$U_{i \rightarrow f} = \frac{e^2}{2} \delta^3(\mathbf{r}_i - \mathbf{r}_f) \sqrt{\omega_1 \omega_2} \sum_n \left[\frac{(\mathbf{r} \mathbf{e}_2)_{2n} (\mathbf{r} \mathbf{e}_1)_{n1}}{\epsilon_1 - \epsilon_n + \omega_1} + \frac{(\mathbf{r} \mathbf{e}_1)_{2n} (\mathbf{r} \mathbf{e}_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right]. \quad (34.9)$$

According to (28.18) in the differential scattering cross section is

$$d\sigma = 2\pi |U_{i \rightarrow f}|^2 \delta(\epsilon_1 + \omega_1 - \epsilon_2 - \omega_2) \frac{d\mathbf{k}_2}{(2\pi)^3},$$

Inserting (34.9) into this expression, we obtain

$$d\sigma = \omega_1 \omega_2^2 d\omega_2 \left| \sum_n \left[\frac{(\mathbf{Q} \mathbf{e}_2)_{2n} (\mathbf{Q} \mathbf{e}_1)_{n1}}{\epsilon_1 - \epsilon_n + \omega_1} + \frac{(\mathbf{Q} \mathbf{e}_1)_{2n} (\mathbf{Q} \mathbf{e}_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right] \right|^2, \quad (34.10)$$

where $\mathbf{Q} = e \mathbf{r}$ is the dipole moment of the electron.

If we go on to higher terms in a power series in $\omega \mathbf{r}$ then by using expression (30.27) for the expansion of a plane wave in spherical waves, it is easy to obtain the following generalization of (34.10):

$$d\sigma = \frac{2\pi}{\omega_1^2} d\omega_2 \sum_{L_1, L_2, M_1, M_2, l_1, l_2} \sum_{\epsilon_1, \epsilon_2} (e_1, Y_{L_1 M_1}^{(l_1)})(e_2, Y_{L_2 M_2}^{(l_2)}) \times \quad (34.11)$$

$$\times \left| \sum_n \left[\frac{(\mathcal{T}_{L_1 M_1}^{(l_1)})_{2n} (\mathcal{T}_{L_2 M_2}^{(l_2)})_{n1}}{\epsilon_1 - \epsilon_n + \omega_1} + \frac{(\mathcal{T}_{L_2 M_2}^{(l_2)})_{2n} (\mathcal{T}_{L_1 M_1}^{(l_1)})_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right] \right|^2,$$

where $\mathcal{T}_{LM}^{(l)}$ is given by the matrix element for multipole radiation $\underline{U}_{i \rightarrow f}$ which was defined in (30.13) and (30.23), by the relation

$$U_{i \rightarrow f} = \mathcal{T}_{LM}^{(l)} \sqrt{d\omega}.$$

2. Emission of Two Photons. Metastable $2s_1/2$ State in Hydrogen.

The matrix $\mathcal{Y}_S^{(2)}$ of (34.2) can also be applied to the problem of simultaneous emission of two photons. The matrix element corresponding to this process is easily obtained from Equations (34.3) through (34.9) by replacing A_1 by A_2 (or k_1 by $-k_1$). Thus so long as the conditions which allow for our above assumptions are maintained (that is, nonrelativistic energies and photon wavelengths large compared to the dimensions of the system), we will have

$$U_{i \rightarrow f} = \frac{e^2}{2} \sqrt{\omega_1 \omega_2} \sum_n \left\{ \frac{(\mathbf{r} \mathbf{e}_2)_{2n} (\mathbf{r} \mathbf{e}_1)_{n1}}{\epsilon_1 - \epsilon_n - \omega_1} + \frac{(\mathbf{r} \mathbf{e}_1)_{2n} (\mathbf{r} \mathbf{e}_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right\}. \quad (34.12)$$

The probability for emission of two photons with momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ (in the intervals $d\mathbf{k}_1 d\mathbf{k}_2$) is given by

$$d\omega_{\mu, \mu_2} = 2\pi |U_{i \rightarrow f}|^2 \delta(\epsilon_1 - \epsilon_2 - \omega_1 - \omega_2) \frac{d\mathbf{k}_1 d\mathbf{k}_2}{(2\pi)^6},$$

from which (34.12) leads to

$$d\omega_{\mu, \mu_2} = \frac{1}{(2\pi)^3} \omega_1^2 \omega_2^2 d\omega_1 d\omega_2 \left| \sum_n \left\{ \frac{(\mathbf{Q} \mathbf{e}_2)_{2n} (\mathbf{Q} \mathbf{e}_1)_{n1}}{\epsilon_1 - \epsilon_n - \omega_1} + \frac{(\mathbf{Q} \mathbf{e}_1)_{2n} (\mathbf{Q} \mathbf{e}_2)_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right\} \right|^2. \quad (34.13)$$

We have been using potentials corresponding to photons in eigenstates of the momentum and polarization. If we consider eigenstates of the angular momentum and parity, then inserting the potentials of (5.19), (5.21) into (34.1) (these are normalized for a photon energy interval $d\omega$) and restricting ourselves as above to the dipole approximation, use of (30.13) leads to

$$U_{i \rightarrow j} = \frac{\omega_1 \omega_2}{\pi} \sqrt{\omega_1 d \omega_1} \sqrt{\omega_2 d \omega_2} \frac{2}{3} \sum_n \left\{ \frac{(Q_{M_1}^{(0)})_{2n} (Q_{M_2}^{(0)})_{n1}}{\epsilon_1 - \epsilon_n - \omega_1} + \frac{(Q_{M_1}^{(0)})_{2n} (Q_{M_2}^{(0)})_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right\},$$

where M_1 and M_2 are the projections of the photon angular momenta. The probability for radiating two photons with energies ω_1 and $\omega_2 = \epsilon_1 - \epsilon_2$, ω_1 is

$$d\omega_{\omega_1} = 2\pi \sum_{M_1, M_2} |U_{i \rightarrow j}|^2 \delta(\epsilon_1 - \epsilon_2 - \omega_1 - \omega_2) = \omega_1^2 \omega_2^2 d\omega_1 \frac{8}{9\pi} \sum_{M_1, M_2} \left| \sum_n \left\{ \frac{(Q_{M_1}^{(0)})_{2n} (Q_{M_2}^{(0)})_{n1}}{\epsilon_1 - \epsilon_n - \omega_1} + \frac{(Q_{M_1}^{(0)})_{2n} (Q_{M_2}^{(0)})_{n1}}{\epsilon_1 - \epsilon_n - \omega_2} \right\} \right|^2, \quad (34.14)$$

Clearly Equation (34.14) should coincide with (34.13) if the latter is integrated over angles and summed over photon polarizations.

The probability for radiating two photons is ordinarily very small compared with that for radiating one photon with an energy $\omega = \omega_1 + \omega_2$. Therefore the first process plays a significant role only for those cases in which single-photon radiation is forbidden by the selection rules. Most important is that case in which the angular momenta of the initial and final states of an atom or nucleus are zero. Then the probability for emission of one photon vanishes identically, since as was shown in Section 4, there exists no single photon state with angular momentum zero.

Since the electron spin is one-half, a state with angular momentum zero is possible only for an even number of electrons (this same statement is true for nucleons). For a single electron, transitions between states with orbital angular momentum zero are also strongly forbidden.

A typical example in which two-photon emission is significantly more probable than single-photon emission is the $2S_{1/2}$ state of hydrogen ($n=2, l=0$). Transition from this state to the ground state $1S_{1/2}$ (with $n=1, l=0$) may be achieved by emission of a photon in the magnetic dipole state ($L=1$, positive parity). In the nonrelativistic approximation the transition probability as calculated by (30.24) vanishes in view of the orthogonality of the radial wave functions for different principal quantum numbers. Nonzero transition probability is obtained only when using the exact relativistic wave functions, and it then turns out to be of the order of 10^{-7} sec^{-1} .

The probability for the $2S_{1/2} \rightarrow 1S_{1/2}$ transition with emission of two photons can be calculated with the aid of (34.14). The dipole moment has matrix elements connecting the s-state only with p-states ($l=1$). Therefore, the sum over s-states in (34.14) applies to principal quantum numbers from $n'=2$ to $n'=\infty$. Each value of n corresponds to three p-states with different magnetic quantum numbers m and the same energy ϵ_n . The sum over n in (34.14) we shall now replace by a double sum $\sum_n \sum_m$, and shall give the matrix elements of the dipole moment two pairs of indices (for the initial state $n=2, \underline{m}=0$, and for the final one $n'=1, \underline{m}=0$). Using the definition (30.12) of the electric dipole moment

$$Q_{M_1}^{(0)} = e' \sqrt{\frac{4\pi}{3}} r Y_{1M_1}^{(0)} \left(e' = \frac{e}{\sqrt{4\pi}} \right)$$

and writing the electron wave functions in the form of products of angular and radial functions

$$\psi_1 = \frac{1}{\sqrt{4\pi}} R_{10}(r), \quad \psi_2 = \frac{1}{\sqrt{4\pi}} R_{20}(r), \\ \psi_{nm} = R_{np}(r) Y_{nm},$$

we can write the matrix elements of the dipole moment in the form

$$(Q_{M_1}^{(0)})_{20}^{e'} = \frac{e'}{\sqrt{3}} Q_{20} \delta_{M_1, m}, \quad (34.15)$$

where

$$Q_{20} = \int_0^\infty R_{20} R_{10} r^3 dr.$$

Similarly

$$(Q_{M_1}^{(0)})_{10}^{e'} = \frac{e'}{\sqrt{3}} Q_{10} (-1)^m \delta_{M_1, -m}.$$

Thus

$$\sum_n (Q_{M_1}^{(0)})_{20}^{e'} (Q_{M_2}^{(0)})_{n1}^{e'} = \sum_n (Q_{M_1}^{(0)})_{20}^{e'} (Q_{M_2}^{(0)})_{n1}^{e'} = \frac{e'^2}{3} Q_{20} Q_{n1} \delta_{M_1, -M_2} (-1)^{M_1},$$

and

$$d\omega_{\omega_1} = \frac{8\pi^2}{27\pi} \omega_1^2 \omega_2^2 d\omega_1 \left| \sum_{n=2}^\infty Q_{20} Q_{n1} \left(\frac{1}{\epsilon_1 - \epsilon_n - \omega_2} + \frac{1}{\epsilon_1 - \epsilon_n - \omega_1} \right) \right|^2. \quad (34.16)$$

The problem now reduces to calculating the sum over principal quantum numbers in (34.16). The radial integrals in (34.15) are generally simple to perform. The summation is performed numerically. The result for the total probability is ¹⁾

$$w = \int_0^{\epsilon_1 - \epsilon_2} \frac{d\omega_{\omega_1}}{d\omega_1} d\omega \approx 7 \text{ sec}^{-1}$$

(or more exactly, $6.5 < w < 8.7 \text{ sec}^{-1}$).

The long lifetime of hydrogen in the metastable $2S_{1/2}$ state was used in the experimental observation of the radiative level shifts.²⁾

¹⁾ G. Breit and E. Teller, *Astrophys. J.* 91, 215 (1940).

²⁾ W. Lamb and R. Retherford, *Phys. Rev.* 72, 241 (1943).

3. Resonance Scattering

Let us now return to the dispersion equation (34.9).

Let us consider a case in which the initial and final electron state is the ground state (with lowest energy)

$$\omega_1 = \omega_0, \quad \epsilon_1 = \epsilon_0.$$

A sum over all excited states enters into (34.9). Let the photon frequency ω_1 be equal to the energy difference between one of the excited states and the ground state of the electron, i.e., $\omega_1 = \epsilon_n - \epsilon_1$ (resonance). Then the matrix element becomes infinite, which indicates that the formula obtained is not valid for the resonance case.

We shall not go into a rigorous theory of this case¹⁾, but shall restrict ourselves merely to a description of the physical reason for the inapplicability of (34.9) close to resonance, and shall indicate an approximate method for treating this region.

The reason for the inapplicability of (34.9) close to resonance is that we have treated the ψ_n as stationary-state wave functions, whose time dependence is of the form $e^{-i\epsilon_n t}$. But excited states are only approximately stationary. Because of the interaction with the electromagnetic field, there exists a finite probability for transition to the ground state. This probability has been determined in Section 30.

An approximately stationary state may be described²⁾ as a state having a "complex energy", whose wave functions have a time dependence of the form

$$\exp \left[-i \left(\epsilon_n - \frac{i}{2} \Gamma_n \right) t \right].$$

The probability for finding the electron in an excited state, which is proportional to $|\psi_n|^2$, now decreases exponentially due to the factor $e^{-\Gamma_n t/2}$. This shows that

$$\Gamma_n = w_n, \quad (34.17)$$

where w_n is the emission probability, as defined by (30.9), and (30.24).

At frequencies close to resonance we can drop all terms in (34.12) other than the resonance term, and in the latter we can replace ϵ_n by $\epsilon_n - i\Gamma_n/2$. Thus, instead of (34.9), we obtain

$$U_{1 \rightarrow f} = -\frac{e^2}{2} \sqrt{\omega_1 \omega_2} \frac{(re)_{1n} (re)_{n1}}{\epsilon_n - \epsilon_1 - \omega_1 - i\Gamma_n/2} \quad (34.17')$$

and the differential scattering cross section becomes

$$d\sigma = \frac{2\pi}{\omega_1} |e_1 Y_{LM}^{(0)*}|^2 |e_2 Y_{LM}^{(0)}|^2 \frac{|(re)_{1n} (re)_{n1}|^2}{(\epsilon_n - \epsilon_1 - \omega_1)^2 + \Gamma_n^2/4} d\Omega_2. \quad (34.18)$$

Similarly, in the general case of a resonance level whose angular momentum and parity differ arbitrarily from that of the ground state, we obtain, according to (34.11),

$$d\sigma = \frac{2\pi}{\omega_1} |e_1 Y_{LM}^{(0)*}|^2 |e_2 Y_{LM}^{(0)}|^2 \frac{|(re)_{1n}^{(0)*} (re)_{n1}^{(0)}|^2}{(\epsilon_n - \epsilon_1 - \omega_1)^2 + \Gamma_n^2/4} d\Omega_2. \quad (34.19)$$

The phenomenon of resonance scattering is essentially the same as that of photon absorption, which was considered in Section 30. In fact, in absorbing a photon, an electron goes into an excited state. This state is not stationary, and the electron returns to the ground state by emitting a photon³⁾. In Section 30, we did not use the cross section to characterize absorption, but the probability for excitation of the atom in a given photon flux. Let us assume that the photons are homogeneously distributed in a frequency region large compared with Γ , and close to the resonance frequency $\omega_1 = \epsilon_n - \epsilon_1$. Then the probability for resonance scattering, that is, the probability for absorption and re-emission of a photon in the frequency interval $d\omega_2$, is

$$dw_{w_2} = \frac{2\pi |(re)_{1n}^{(0)*} (re)_{n1}^{(0)}|^2}{(\epsilon_n - \epsilon_1 - \omega_1)^2 + \Gamma_n^2/4} d\omega_2 \quad (\omega_2 = \omega_1). \quad (34.20)$$

Let us integrate this probability over frequencies. Since the region of integration may be said to have a width of order $\Gamma \ll \omega_1$, the matrix element can be taken out from under the integral sign. In the remaining integral, the limits may be considered infinite, so that we write

$$\int_{-\infty}^{\infty} \frac{d\omega_1}{(\epsilon_n - \epsilon_1 - \omega_1)^2 + (\Gamma_n/2)^2} = \frac{2\pi}{\Gamma_n}.$$

Thus the probability of scattering,

$$w = 2\pi |(re)_{1n}^{(0)*}|^2 \frac{2\pi |(re)_{n1}^{(0)}|^2}{\Gamma_n}, \quad (34.21)$$

is the same as the absorption probability, since Γ_n is the probability for transition from the excited state to the ground state, given by

$$\Gamma_n = 2\pi |(re)_{1n}^{(0)}|^2.$$

¹⁾ We assume here that between the ground state and the resonance level there is no other level for which the transition probability is comparable to that for the ground state. If this were not so, there would take place scattering for which the scattered frequency $\omega_2 \neq \omega_1$.

²⁾ The systematic treatment of the resonance scattering using the general theory of radiative corrections will be given in Section 45, paragraph 3.

³⁾ See, for example, L. Landau and E. Lifshits, Quantum Mechanics (State Tech. Press, 1948) Section 119.

A similar method can be used for treating the general case of resonance scattering, as well as two-photon emission. The latter process reduces to "cascade" transition from the excited state to the ground state by means of an intermediate level. The corresponding matrix element is

$$U_{i \rightarrow f} = \frac{(T_{L,M}^{\lambda})_{2n} (T_{L,M}^{\lambda})_{n1}}{\epsilon_1 - \epsilon_n - \omega_1 + i\Gamma/2}, \quad (34.22)$$

and the emission probability is

$$dw = 2\pi \frac{|(T_{L,M}^{\lambda})_{2n}|^2 |(T_{L,M}^{\lambda})_{n1}|^2}{(\epsilon_1 - \epsilon_n - \omega_1)^2 + (\Gamma/2)^2} d\omega_1. \quad (34.23)$$

In analogy with (34.21), the total probability is the same as that for transition from state 1 to state $\underline{n}$ (followed by transition to state 2).

We have not taken into account the degeneracy of the intermediate state $\underline{n}$. If the angular momentum of this state is $\underline{j}_n$, the matrix element contains a sum over states with different $\underline{m}$.

$$U_{i \rightarrow f} = \frac{1}{\epsilon_1 - \epsilon_n - \omega_1 + i\Gamma/2} \sum_{\underline{m}} (T_{L,M}^{\lambda})_{2n;\underline{m}} (T_{L,M}^{\lambda})_{n1;\underline{m}}.$$

For fixed $\underline{m}_1$ and $\underline{m}_2$ in the initial and final electron states and for fixed values of $\underline{M}_1$ and $\underline{M}_2$ for the photons, however, the value of $\underline{m}$ is uniquely determined by the selection rules

$$m = m_1 - M_1 = m_2 + M_2. \quad (34.24)$$

The total transition probability, averaged over the magnetic quantum numbers of the initial state and summed over those of the final electron state and of the two photons, is

$$dw = \frac{1}{2j_1 + 1} \sum_{M_1, M_2} \sum_{m_1, m_2} dw_{\underline{m}_1, \underline{m}_2}^{\underline{m}_1, \underline{m}_2} = \frac{1}{2j_1 + 1} \sum_{\underline{m}} \sum_{M_1, M_2} dw_{\underline{m}, \underline{m}}^{\underline{m}, \underline{m}}, \quad (34.25)$$

$$m_1 = m + M_1, \\ m_2 = m - M_2,$$

where $dw_{\underline{m}_1, \underline{m}_2}^{\underline{m}_1, \underline{m}_2}$ is the probability given by (34.23).

4. Angular Correlation in Successive Emission of Two Photons.

The angular distribution for the emission probability of a photon with given angular momentum $\underline{L}$ and given projections of this angular momentum $\underline{M}$ is determined by (4.25). Let us denote this distribution function by F_{LM} , writing

$$F_{LM}^{(n)} = |Y_{LM}^{(n)}|^2, \quad (34.26)$$

where $\underline{n}$ is the directional unit vector. Due to the degeneracy related to the spatial symmetry of free atoms and nuclei, emission of photons with various values of $\underline{m}$ is equally probable. Therefore the emission probability averaged over the projections of the angular momentum in the initial and final states is isotropic, that is

$$\sum_{\underline{M}} F_{LM} = \text{const.}$$

For cascade emission of two photons, $\underline{M}_1$ and $\underline{M}_2$ are related by Equation (34.24). Therefore there is a correlation between the directions of the emitted photons. Let us find this correlation function.

Let $f_{\underline{M}_1 \underline{M}_2}$ be the probability for emission of photons with given $\underline{M}_1$ and $\underline{M}_2$ into the corresponding elements of solid angle. Since we are interested only in angular correlation, we shall drop factors independent of the directions $\underline{n}_1, \underline{n}_2$ and independent of the $\underline{M}_1, \underline{M}_2$ (the resulting constant can be found from the normalization condition).

It follows from (34.23) that

$$f_{\underline{M}_1 \underline{M}_2} = |(T_{L,M_1}^{\lambda})_{2n;\underline{m}} (T_{L,M_2}^{\lambda})_{n1;\underline{m}}|^2 F_{L,M_1}(\underline{n}_1) F_{L,M_2}(\underline{n}_2),$$

where $\underline{j}_1, \underline{j}_2, \underline{j}$, and $\underline{m}_1, \underline{m}_2, \underline{m}$ are the angular momenta and their projections for the initial, intermediate, and final electron states. In order to average over all orientations of the angular momenta this equation must be summed over $\underline{M}_1, \underline{M}_2$ and $\underline{m}$. In doing this we may choose the $\underline{z}$ axis along the direction of the first photon. Then the correlation function $F(\theta)$, where θ is the angle between the two photons, can be written

$$F(\theta) = \sum_{M_1, M_2, m} |(T_{L,M_1}^{\lambda})_{2n;\underline{m}} (T_{L,M_2}^{\lambda})_{n1;\underline{m}}|^2 F_{L,M_1}(\underline{n}_0) F_{L,M_2}(\theta) \quad (34.27)$$

(where $\underline{n}_0$ is directed along the $\underline{z}$ axis).

According to (34.26) and (4.5) F_{LM} can be written in terms of spherical functions. Since $Y_{LM}^{(0)}(0)$ is nonzero only when $\underline{m} = 0$, then as can be easily shown on the basis of (4.15) (see also the table on p. 29), $\underline{m}$

$$F_{LM}(\underline{n}_0) = 0 \text{ for } |M| \neq L, \\ F_{LM}(\underline{n}_0) = F_{L-1}(\underline{n}_0).$$

Thus (34.27) becomes (up to a constant factor)

$$F(0) = \sum_{m=-j}^j \sum_{M_2=-L_2}^{L_2} \sum_{M_1=-L_1}^{L_1} (T_{L_2 M_2}^{L_1 M_1})_{j j} (T_{L_1 M_1}^{L_2 M_2})_{j j} F_{L_2 M_2}(0),$$

$$m_2 = m - M_2, \quad m_1 = m + M_1.$$

We note that we could pick the z axis along the direction of the second photon with equal justification. Then $F(0)$ will contain not $F_{L_2 M_2}$, but $F_{L_1 M_1}$. In practice it is convenient to write $F(0)$ in terms of the distribution $F_{LM}(0)$ of the photon with the lower value of L . The matrix elements $(T_{L_2 M_2}^{L_1 M_1})_{j j}$ and $(T_{L_1 M_1}^{L_2 M_2})_{j j}$ can be found in general. Since the multipole moment operator $Q_{LM}^{(\lambda)}$ is an L -vector,¹⁾ that is, a quantity which transforms according to the $(2L+1)$ -dimensional irreducible representation of the rotation group, and the wave functions $\psi_{j_2 m_2}$, $\psi_{j_1 m_1}$, and $\psi_{j m}$ are also L -vectors (with $L = j_2, j_1, j$), the $(T_{LM}^{L'})_{j j}$ are given, up to an m -independent factor by the coefficients of the series

$$\psi_{j_1 m_1} \psi_{j_2 m_2} = \sum_L C_{LM}^{j_1 m_1 j_2 m_2} \psi_{j m}$$

(see (57.10)). Thus

$$F(0) = \sum_{m=-j}^j \sum_{M_2=-L_2}^{L_2} \sum_{M_1=-L_1}^{L_1} |C_{LM}^{j_1 m_1 j_2 m_2}|^2 F_{LM}(0). \quad (34.28)$$

From the general form of $F_{LM}(0)$ it is easy to see that the correlation function $F(\theta)$ can be written as a series of even Legendre polynomials,

$$F(\theta) = \sum_{N=0}^N A_N P_N(\cos \theta), \quad (34.29)$$

where N is the smallest of the three numbers $2j_1$, $2j_2$ and $2j$:

$$N \leq 2j_1, \quad N \leq 2j_2, \quad N \leq 2j.$$

Thus there is no angular correlation if j , the angular momentum of the intermediate state, is zero or one-half.

The summation of (34.28) can be carried out²⁾ if we make use of a theorem for the C coefficients [see (57.7)]. We shall merely present here the values of A_N of Equation (34.29) obtained in this way³⁾:

¹⁾ See Section 57.

²⁾ W. Gardner, *Proc. Phys. Soc.*, B4, 2, 16 (1951).

³⁾ S. Lloyd, *Phys. Rev.*, 83, 71 (1951).

$$A_{2n} = (2n+1) b_{2n}(L_1) b_{2n}(L_2) t_{2n}, \quad (34.30)$$

where

$$b_k(L) = \left[1 - \frac{k(k+1)}{2L(L+1)} \right] \frac{(2L+1)! k! \left(L + \frac{k}{2} \right)!}{(2L+k+1)! \left(\frac{k}{2} \right)! \left(L - \frac{k}{2} \right)!},$$

and

$$t_k = 1 \quad \text{for } j_1 = j \pm L_1, \quad j_2 = j \mp L_2;$$

$$t_k = \frac{(2j-k)!(2j+k+1)!}{(2j)!(2j+1)!} \quad \text{for } j_1 = j - L_1, \quad j_2 = j - L_2;$$

$$t_k = \frac{(2j)!(2j+1)!}{(2j-k)!(2j+k+1)!} \quad \text{for } j_1 = j + L_1, \quad j_2 = j + L_2;$$

$$t_k = -\frac{2j+3}{j} \quad \text{for } j_1 = 1, \quad j_2 = j - L_2;$$

$$t_k = -\frac{2j-1}{j+1} \quad \text{for } j_1 = 1, \quad j_2 = j + L_2;$$

$$t_k = \frac{(2j-1)(2j+3)}{j(j+1)} \quad \text{for } j_1 = j_2 = 1.$$

Angular correlation of photons takes place during γ -emission by nuclei. In order to observe the correlation it is necessary that the transition probability from the intermediate state to the ground state be sufficiently large (i.e., that the intermediate state be "short-lived"). Otherwise, the correlation will be destroyed by interaction between the nucleus and the electron shells. This interaction (which is responsible for the hyperfine structure of atomic levels) causes the orientation of the nuclear angular momentum to change in the intermediate state, so that the quantum number m no longer satisfies (34.23). The condition that the angular correlation not be destroyed reduces to

$$\Gamma_n \gg \nu,$$

where ν is the magnitude of the hyperfine splitting.

Note. The theory of bremsstrahlung developed in this Chapter refers to the collision of electrons or photons with isolated atoms. At high energies, however, as has been shown by Landau and Pomeranchuk⁴⁾, when the finite density of matter is taken into account, the theory is altered greatly. We present here the fundamental results of this work.

It can be shown that when emission takes place, distances along the direction of motion of the electron $\sim l = \epsilon_1^2 / m\omega$ (the notation is that of Section 31) are of importance. On the other hand, the effective energy loss to radiation takes place over a distance $L = \Phi / N$ (the so-called "cascade unit of length"), where Φ is given by (31.24), and N is the density of atoms. The theory of bremsstrahlung on an isolated atom becomes applicable when $L \gg l$, that is, when $\epsilon_1 \gg \sqrt{m^2 \omega L}$. If we take account of multiple elastic scattering, then this limit is reduced, and the theory becomes applicable when $\epsilon_1 \gg \sqrt{m^2 \omega L} \frac{N}{E}$, where $E = m \sqrt{4\pi \cdot 137} = 21$ Mev.

This is essentially true for arbitrary ϵ_1 for sufficiently small ω . It is found that as $\omega \rightarrow 0$, the photon spectrum

⁴⁾ L. Landau and I. Pomeranchuk, *Proc. Acad. Sci. USSR* 92, 3 (1953).

is of the form $d\omega/\omega$. Instead of $d\omega/\omega$, which means that the infrared catastrophe does not occur. In general when $\epsilon_2 = \epsilon_1 - \omega < E_0 = m^2 L^2 / E_1^2$ the ordinary formulas are valid, and when $\epsilon_2 \gg E_0$, the matrix elements are multiplied by an additional factor or order $\sqrt{E_0/\epsilon_1 \epsilon_2}$. In the ordinary theory the energy losses per unit length are proportional to ϵ_1 . When $\epsilon_1 \gg E_0$, they are proportional to $\sqrt{\epsilon_1}$. The mean free path for photons with respect to pair creation, for energies $\omega \gg E_0$, increases as $\sqrt{\omega}$. Thus at very high energies electrons and photons become much more penetrating.

CHAPTER VII

RETARDED INTERACTION BETWEEN TWO CHARGES

§ 35. Interaction Matrix of Two Charges. Retarded Potentials.

1. Interaction Matrix of Two Charges.

In this chapter we shall consider those interactions of two charges which do not involve emission or absorption of photons. We shall, therefore, concern ourselves only with those elements of the S matrix which connect initial and final states of the following types. The initial state contains two electrons¹⁾ in definite individual states, and there are no photons. The final state contains two electrons in different states and also contains no photons.

We shall restrict our considerations of such processes to the first nonvanishing approximation. Since the first order S matrix contains only terms corresponding to emission or absorption of a photon²⁾, we must consider the second-order matrix $S^{(2)}$.

According to (20.15) the general expression for $S^{(2)}$ can be written

$$S^{(2)} = -\frac{1}{2} \int P \{ j_i(1) j_k(2) \} P \{ A_i(1) A_k(2) \} d^4 x_1 d^4 x_2, \quad (35.1)$$

where P is the chronological operator, j is the current operator, and A is the potential operator; the arguments 1 and 2 denote the appropriate four-dimensional points x_1 and x_2 .

We shall denote the set of elements of this matrix which connect states with no photons by the symbol $S_S^{(2)}$. We arrive at an expression for $S_S^{(2)}$ by replacing the operator $P \{ A_i(1) A_j(2) \}$ in (35.1) by its vacuum expectation value $\langle P \{ A_i(1) A_j(2) \} \rangle_0$. According to (16.25),

$$\langle P \{ A_i(1) A_j(2) \} \rangle_0 = \frac{1}{2} \delta_{ij} D^F(1, 2),$$

¹⁾ We are considering here electrons with both signs of the charge, that is, electrons and positrons. The general formulas of this section will be valid for interactions between charges both of like signs and of opposite signs.

²⁾ We shall consider the external field included in the electron Hamiltonian. Thus, an individual electron state will in general be thought of as an electron state in an external field.

where $\underline{D}^F$ is defined by (16.29) and (16.31).

The operator $j_i(1)j_k(2)$ is a product of four electron creation or annihilation operators. As was mentioned above, we shall be concerned with those matrix elements of $S^{(2)}$ which give two electrons both in the initial and final states. Therefore, nonzero results are obtained only from those terms of $j_i(1)j_k(2)$ which contain two creation and two annihilation operators, where all four refer to different individual electron states. Therefore, all four factors anticommute and they commute in pairs, so that

$$j_i(1)j_k(2) = j_k(2)j_i(1),$$

and the operator P in (35.1) can be dropped.

Thus, the scattering matrix for two charges can be written

$$eS^{(2)} = -\frac{1}{4} \int \int j_i(2) D^F(2, 1) j_k(1) d^4x_2 d^4x_1. \quad (35.2)$$

Let us point out the following features of the structure of expression (35.2). In nonrelativistic quantum mechanics the interaction of particles is described by the potential energy of their interaction $U(|\mathbf{r}_1 - \mathbf{r}_2|)$, which is a function of the separation between the particles. The potential energy operator U (in the occupation-number representation) can be written

$$U = \frac{1}{2} \int \rho(\mathbf{r}_1) U(\mathbf{r}_1, \mathbf{r}_2) \rho(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2,$$

where $\rho = \psi^* \psi$ is the density operator.

In electrodynamics there exists no potential energy of interaction which depends on the coordinates of the interacting particles at the same instant of time. There exists only a point-interaction, in that electrons interact not directly, but through the electromagnetic field. The interaction energy operator is replaced by the matrix $eS^{(2)}$ which plays a similar role, and is given in the first approximation by (35.2). The integrand in (35.2) contains the coordinates (including, however, the time coordinate) only of the electrons, but not of the electromagnetic field. Therefore,

$$\frac{e^2}{2} D^F(2, 1)$$

can be called the interaction function of two charges.

We shall see later (see Section 37) that in the nonrelativistic approximation this function will indeed lead to the expression for the interaction energy between two charges (in particular, the Coulomb interaction).

2. General Form of the Matrix Element.

Let us find a general expression for the matrix elements of $eS^{(2)}$. In particular, let us consider the case in which both particles are electrons. The formulas so obtained can easily be rewritten to include the case of two positrons or an electron and a positron.

Let $\underline{A}$ and $\underline{B}$ be the set of quantum numbers of the initial electron states, and $\underline{C}$ and $\underline{D}$ those of the final states. Then those terms of

$$j_i(1)j_k(2) = e^2 (\bar{\psi}(1) \gamma_i \psi(1)) (\bar{\psi}(2) \gamma_k \psi(2))$$

which give nonzero contributions are

$$\begin{aligned} \frac{1}{2!} j_i(1)j_k(2) \rightarrow & a_C^* a_A^* a_B^* a_D (\bar{\psi}_C(1) \gamma_i \psi_A(1)) (\bar{\psi}_D(2) \gamma_k \psi_B(2)) + \\ & + a_D^* a_B^* a_A^* a_C (\bar{\psi}_C(2) \gamma_k \psi_B(2)) (\bar{\psi}_D(1) \gamma_i \psi_A(1)) + \\ & + a_C^* a_B^* a_A^* a_D (\bar{\psi}_C(1) \gamma_i \psi_B(1)) (\bar{\psi}_D(2) \gamma_k \psi_A(2)) + \\ & + a_D^* a_A^* a_B^* a_C (\bar{\psi}_C(2) \gamma_k \psi_A(2)) (\bar{\psi}_D(1) \gamma_i \psi_B(1)), \end{aligned} \quad (35.3)$$

where a^* and a are creation and annihilation operators for electrons in the corresponding states.

If we make use of the anticommutation of all the operators in this expression, we can order them in the form of the normal product

$$a_C^* a_D^* a_A^* a_B$$

whose inatrix element is unity.

Now the operator products in (35.3) can be factored out, and the first two terms then enter with a + sign, and the second two terms with a - sign. We note further, that the first term of (35.3) differs from the second only in the transposition of the arguments (1) and (2). When inserted into (35.2) these give identical results, since D^F is symmetric with respect to interchange of (1) and (2). The third and fourth terms in (35.3) also differ by the transposition of their arguments.

Thus, the matrix element $\frac{e^2}{2} \frac{1}{2!} \rightarrow \frac{1}{2!}$ can be written

$$eS_{i \rightarrow f}^{(2)} = -\frac{e^2}{2} \int (\bar{\psi}_D(2) \gamma_k \psi_B(2)) D^F(2, 1) (\bar{\psi}_C(1) \gamma_i \psi_A(1)) - \\ - \bar{\psi}_C(2) \gamma_i \psi_B(2) D^F(2, 1) \bar{\psi}_D(1) \gamma_k \psi_A(1) d^4x_2 d^4x_1. \quad (35.4)$$

According to the general rules for graphical representation of the matrix elements, the two terms of (35.4) correspond to the two diagrams shown in Fig. 47. They differ only in that the indices of the final states are interchanged. Therefore, the second term in (35.4) (and the second diagram of Fig. 47) is called the exchange term of the first. The - sign of the exchange term in (35.4) corresponds to the antisymmetry of the two-electron wave function in configuration space.

We shall write the matrix element (35.4) in the form

$$S_{AB;CD}^{(2)} = S_{AB;CD}^{(2)} - S_{AB;DC}^{(2)}, \quad (35.5)$$

where

$$S_{AB;CD}^{(2)} = -\frac{e^2}{2} \int \bar{\psi}_B(2) \gamma_i \psi_B(2) D^F(2,1) \bar{\psi}_A(1) \gamma_i \psi_A(1) d^4x_2 d^4x_1, \quad (35.6)$$

and $S_{AB;DC}^{(2)}$ differs from $S_{AB;CD}^{(2)}$ by the interchange of C and D .

Expression (35.4) can be used also for the case of two positrons if (as opposed to the case of two electrons) C and D denote the initial states, and A and B the final states. The functions ψ_A^* , ψ_B^* , ψ_C^* , ψ_D^* in (35.4) then denote not positron wave functions, but negative-frequency wave functions corresponding to them.

Indeed, since the operator ψ contains positron creation operators, and $\bar{\psi}$ contains positron annihilation operators, in order to go to the case of positrons, a^* should be replaced by b and a by b^* in (35.3).

Bringing the operators into normal form

$$b_C b_A^* b_D b_B^* = b_A^* b_B^* b_D b_C$$

and making use of the fact that the matrix element of this operator is unity, we arrive at (35.4).

If one of the particles is an electron and the other is a positron, we shall use the following notation: A is the initial electron state, C is the final electron state, B is the final positron state, and D is the initial positron state. In expression (35.3) we must then replace

$$a_D^* \text{ by } b_D^* \text{ and } a_B \text{ by } b_B^*.$$

Again, writing the product of a and b operators in normal form,

$$a_A^* a_B b_D b_B^* = -a_A^* b_B^* b_D b_A,$$

we obtain

$$S_{AB;CD}^{(2)} = -(S_{AB;CD}^{(2)} - S_{AB;DC}^{(2)}). \quad (35.7)$$

Equation (35.7) differs from (35.5) only in its sign, which corresponds to the fact that the electron and positron have opposite charges. We note that the exchange term in (35.7) (the second diagram in Fig. 47) now corresponds to virtual annihilation and subsequent creation of the electron-positron pair.

We shall apply the fundamental equation (35.2) also to the interaction of an electron with a proton. In this case $j(1)$ can denote the electron current, and $j(2)$ the proton current, or vice versa. Let states A and C be electron states, and B and D be proton states. Then the matrix elements of $j_1^{(1)} j_1^{(2)}$ will not contain exchange terms, i.e., instead of (35.3) we will have

$$j_1^{(1)} j_1^{(2)} = 2e^2 a_D^* a_B^* a_A \bar{\psi}_C(1) \gamma_i \psi_A^*(1) (\bar{\psi}_B(2) \gamma_i \psi_B(2)).$$

We will then obtain

$$S_{AB;CD}^{(2)} = -S_{AB;CD}^{(2)}, \quad (35.8)$$

where $S_{AB;CD}^{(2)}$ is given by (35.6). [The - sign in (35.8) corresponds to the fact that the electron and proton have opposite charges.] The matrix element of (35.8) is represented by the first of the diagrams of Fig. 47.

3. Retarded Potentials and Transition Currents.

The matrix element $S_{AB;CD}^{(2)}$ can be written in a form which is very convenient for applications and can be given a simple physical explanation. Let us transform Equation (35.6) by separating out the time-dependent factors of the wave functions. Let

$$\psi_A(1) = \psi_A(r_1) e^{-i\omega_A t_1}$$

with similar equations for ψ_C , ψ_B , and ψ_D . Further, let us write $D^F(2,1)$ in the form of (16.33), namely

$$\frac{1}{2} D^F(2,1) = -\frac{i}{(2\pi)^4} \int \frac{e^{ik(r_1-r_2)-i\omega(t_1-t_2)}}{k^2 - \omega^2} d^4k d\omega. \quad (35.9)$$

We then obtain

$$S_{AB;CD}^{(2)} = \frac{ie^2}{(2\pi)^4} \int \bar{\psi}_A(r_1) \gamma_i \psi_A(r_1) dr_1 \int \bar{\psi}_B(r_2) \gamma_i \psi_B(r_2) dr_2 \times \\ \times \int e^{ik(r_1-r_2)} dk \int \frac{d\omega}{k^2-\omega^2} \int e^{i(\omega_0-\omega_A-\omega)t_1} dt_1 \int e^{i(\omega_B-\omega_D+\omega)t_2} dt_2. \quad (35.10)$$

Let us first perform the integration over t_2 :

$$\int e^{i(\omega_D-\omega_B+\omega)t_2} dt_2 = 2\pi\delta(\omega_D-\omega_B+\omega).$$

Integration over ω now reduces to replacing ω by

$$\omega_{BD} = \omega_B - \omega_D.$$

Now let us integrate over k . Integration over angles gives

$$\int \frac{e^{ik(r_1-r_2)}}{k^2-\omega^2} dk = \frac{4\pi}{|r_1-r_2|} \int_0^\infty \frac{\sin x |r_1-r_2|}{x^2-\omega^2} x dx.$$

Integration over k must be performed according to the rules of Sections 16 for dealing with the poles of the integrand, i.e., by replacing k^2 by $k^2 - i0$; this gives

$$\int_0^\infty \frac{\sin x |r_1-r_2|}{x^2-\omega^2} x dx = \frac{\pi}{2} e^{i\pi} |r_1-r_2|.$$

Inserting this expression into (35.10) and introducing the "transition currents"

$$\left. \begin{aligned} j_{CA}(1) &= e\bar{\psi}_C(1)\gamma\psi_A(1), & j_{CA}(r_1) &= e\bar{\psi}_C(r_1)\gamma\psi_A(r_1), \\ j_{DB}(2) &= e\bar{\psi}_D(2)\gamma\psi_B(2), & j_{DB}(r_2) &= e\bar{\psi}_D(r_2)\gamma\psi_B(r_2), \\ j_{CA}(r_1, t) &= j_{CA}(r_1) e^{i\omega_{CA}t}, \\ j_{DB}(r_2, t) &= j_{DB}(r_2) e^{i\omega_{DB}t}, \end{aligned} \right\} \quad (35.11)$$

we obtain the following expression for the matrix element $S_{AB;CD}^{(2)}$:

$$S_{AB;CD}^{(2)} = \frac{i}{4\pi} \int j_{CA}(r_1, t) j_{DB}(r_2, t) \frac{e^{i(t_1-\omega_{BD}|r_1-r_2|)}}{|r_1-r_2|} dr_1 dr_2 dt. \quad (35.12)$$

We note the integral over time in (35.12) is

$$2\pi\delta(\omega_{CA}+\omega_{DB}), \quad (\omega_{CA}=\omega_C-\omega_A, \quad \omega_{DB}=\omega_D-\omega_B).$$

The δ -function which appears here expresses the conservation of energy. The frequencies of the transition currents are equal in magnitude and of opposite sign. Let us assume, specifically, that

$$\omega_A > \omega_C,$$

so that

$$\omega_B < \omega_D$$

and

$$\omega_{CA} = -\omega_{DB} > 0.$$

With this condition we may rewrite (35.12) in the form

$$S_{AB;CD}^{(2)} = i \int dt \int j_{DB}(r_2, t) A_{CA}(r_1, t) dr_2 = ie \int \bar{\psi}_D(2) \hat{A}_{CA}(2) \psi_B(2) d^4x_2, \quad (35.13)$$

where

$$A_{CA}(r_1, t) = \frac{1}{4\pi} \int j_{CA}(r_1) \frac{e^{-i\omega_{CA}t' + i\omega_{CA}0^+|r_1-r_2|}}{|r_1-r_2|} dr_1,$$

or

$$A_{nA}(r_2, 0) = \frac{1}{4\pi} \int \frac{j_{nA}(r_1, t - |r_1 - r_2|)}{|r_1 - r_2|} dr_1, \quad (35.14)$$

It is remarkable that (35.13) is the same as the expression for the matrix element of the first-order matrix $S^{(1)}$ for transition for an electron from a state $\underline{B}$ to a state $\underline{D}$ due to an external field whose potential is $\underline{A}_{CA}$. Here $\underline{A}_{CA}$ is determined by the transition current $\underline{j}_{CA}$ of the other particle from state $\underline{A}$ to state $\underline{C}$ in the same way as in classical electrodynamics, i.e., by the retarded potential Equation (35.14). Of course, the differential equations for the classical electrodynamic potentials are valid here also:

$$\square A_{CA} = -j_{CA},$$

$$\frac{\partial(A_{CA})}{\partial x_i} = 0.$$

We note that Equation (35.12) for $S_{AB; CD}^{(2)}$ is symmetric with respect to the currents. On going from that equation to (35.13) we have separated these into the current $\underline{j}_{DB}$ which "perceives" the external field $\underline{A}_{CA}$, and the current $\underline{j}_{CA}$ which generates this field. The current here chosen to generate the field is that with positive frequency ($\omega_{AC} > 0$ by assumption; the transition is from a higher energy state to a lower one). If we consider the current with negative frequency the generating one, i.e., if we write (35.12) in the form

$$S_{AB; CD}^{(2)} = i \int \int \partial_{\lambda\lambda} A_{DB}(1) d^4x_1,$$

then instead of (35.14) we obtain $\underline{A}_{DB}$ in the form of an "advanced" potential

$$A_{DB}(r_2, 0) = \frac{1}{4\pi} \int \frac{j_{DB}(r_1, t + |r_1 - r_2|)}{|r_1 - r_2|} dr_1.$$

Let us now go over from the S matrix to the matrix of the effective perturbation energy for a system of two charges, as defined by (28.17), namely

$$S_{i \rightarrow f} = \exp \left[-2\pi i \int_{t_i}^{t_f} \delta(\omega_A + \omega_B - \omega_C - \omega_D) \right]$$

which in our case becomes

$$S_{AB; CD} = -2\pi i U_{AB; CD} \delta(\omega_{AC} + \omega_{BD}),$$

Integrating over t in (35.12), we obtain

$$U_{AB; CD} = -\frac{1}{4\pi} \int \int j_{nA}(r_1) j_{DB}(r_2) \frac{e^{i(\omega_{AC} + \omega_{BD})(|r_1 - r_2|)}}{|r_1 - r_2|} dr_1 dr_2. \quad (35.15)$$

Expression (35.15) can also be written

$$U_{AB; CD} = - \int j_{DB}(r_2) A_{CA}(r_2) dr_2, \quad (35.16)$$

where

$$A_{CA}(r_2) = \frac{1}{4\pi} \int \frac{j_{CA}(r_1)}{|r_1 - r_2|} e^{i\omega_{AC}(|r_1 - r_2|)} dr_1.$$

The potentials $\underline{A}_{CA}$ satisfy the differential equations of classical electrodynamics for monochromatic fields,

$$\left. \begin{aligned} \Delta A_{CA} + \omega_{CA}^2 A_{CA} &= -j_{CA}, \\ \text{div } A_{CA} - \omega_{CA} (A_{CA})_{,0} &= 0. \end{aligned} \right\} \quad (35.17)$$

The results of this section give a rigorous basis to the semiclassical method for treating the interaction of electrons with the electromagnetic field (so-called "correspondence principle"). In this method the quantum properties of the electromagnetic field are not taken into account, and the equations of classical electrodynamics are used with the currents replaced by the transition currents. It can also be shown that the classical energy flux corresponding to the potentials of (35.14) is equal to the probability per unit time for radiating a photon during a transition from state $\underline{A}$ to state $\underline{C}$, multiplied by the energy of the photon $\hbar\omega_{AC}$.

¹ See, for instance, D. Blokhintsev, Elements of Quantum Mechanics (State Tech. Press, 1949) Sections 86, 87.

9. 36. Electron and Positron Scattering by an Electron.

1. Electron-Electron Scattering.

The simplest process that can be investigated with the aid of $\frac{e^4}{(2\pi)^4}$ is collision of two electrons. We shall choose the initial and final electron states to be momentum and polarization eigenstates. In analogy with (29.4) we shall write their wave functions in the form

$$\begin{aligned}\psi_A &= u_1 e^{ip_1 x}, \\ \psi_B &= u_2 e^{ip_2 x}, \\ \psi_C &= u_3 e^{ip_3 x}, \\ \psi_D &= u_4 e^{ip_4 x}.\end{aligned}$$

Here p_1 and p_1' are the initial, and p_2 and p_2' the final four-momenta; the unit amplitudes u_1, u_2, u_3, u_4 correspond to definite polarizations $\mu_1, \mu_2, \mu_1', \mu_2'$. The wave functions are normalized for a unit volume.

Inserting these wave functions and Equation (16.31) for D_F^E into (35.4), and integrating over x_1, x_2 and k , according to the general rules formulated in Section 24 we obtain the following expression for the matrix element:

$$S_{i \rightarrow f} = ie^2 M (2\pi)^4 \delta(p_1 + p_1' - p_2 - p_2'),$$

where

$$M = \frac{(\bar{u}_3 \gamma_\mu u_1)(\bar{u}_4 \gamma_\nu u_2)}{(p_2 - p_1)^2} - \frac{(\bar{u}_3 \gamma_\nu u_1)(\bar{u}_4 \gamma_\mu u_2)}{(p_2 - p_1')^2}. \quad (36.1)$$

Figure 48 shows the same diagrams as those of Fig. 47, in this case, for free electrons.

Now using the general rules of Section 28 we can write the expression for the differential cross section $d\sigma$, namely

$$d\sigma = \frac{e^4}{(2\pi)^4} |M|^2 d p_2 d p_2' \frac{1}{J} \delta(p_1 + p_1' - p_2 - p_2') \delta(\epsilon_1 + \epsilon_1' - \epsilon_2 - \epsilon_2'), \quad (36.2)$$

where J is the particle current density.

In the center-of-mass system of the two electrons ($p_1 + p_1' = p_2 + p_2' = 0$), we have

$$J = v_0,$$

where v_0 is the relative velocity of the particles. Eliminating the δ -function from (36.2), we obtain

$$d\sigma = \frac{e^4}{(2\pi)^4} \left(\frac{p}{v_0} \right)^2 |M|^2 d\Omega, \quad (36.3)$$

where p is the absolute value of the momentum, and $d\Omega$ is the element of solid angle in the direction of scattering (in the center-of-mass system).

The expression for the differential scattering cross section must now be summed over electron spins in the final states and averaged over spins of the initial states. To do this we replace $|M|^2$ in (36.3) by $\frac{1}{4} \Sigma |M|^2$, where Σ denotes summation over μ_1, μ_2, μ_1' and μ_2' . This summation can be performed with the aid of the method described in Section 9 and is not essentially more complicated than previously described, despite the fact that in (36.1) we have not two but four electron states. In analogy with (9.20) we obtain

$$\begin{aligned}\Sigma |M|^2 &= \frac{1}{16\pi^2 v_0^2} \left\{ \frac{1}{(p_2 - p_1)^2} \text{Sp} \gamma_\mu (m - i\hat{p}_1) \gamma_\nu (m - i\hat{p}_2) \times \right. \\ &\quad \times \text{Sp} \gamma_\mu (m - i\hat{p}_1') \gamma_\nu (m - i\hat{p}_2') + \\ &\quad + \frac{1}{(p_2' - p_1)^2} \text{Sp} \gamma_\mu (m - i\hat{p}_1) \gamma_\nu (m - i\hat{p}_2) \text{Sp} \gamma_\mu (m - i\hat{p}_1') \gamma_\nu (m - i\hat{p}_2') - \\ &\quad \left. - \frac{2}{(p_2 - p_1)^2 (p_2' - p_1)^2} \text{Sp} \gamma_\mu (m - i\hat{p}_1) \gamma_\nu (m - i\hat{p}_2) \gamma_\mu (m - i\hat{p}_1') \gamma_\nu (m - i\hat{p}_2') \right\} \quad (36.4)\end{aligned}$$

[$\gamma = \beta \gamma \beta$, see (9.16)].

In these expressions we can replace $\bar{\gamma}_j$ everywhere by γ_j . Indeed, since j is a summation index,

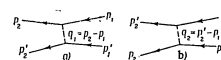


Fig. 48

and since

$$\bar{\gamma} = -\gamma; \quad \bar{\gamma}_4 = \gamma_4$$

we obtain

$$\bar{\gamma}_j \dots \bar{\gamma}_j = \gamma_j \dots \gamma_j$$

By using (28.12) Equation (36.4) is reduced to a sum of terms containing the traces of products of two or four γ matrices, which are given by (9.23).

As a result we obtain the following formula for the differential cross section as a function of energy, velocity, and scattering angle θ in the center-of-mass system:

$$d\sigma = \frac{\alpha^2 \left(\frac{e}{m}\right)^2}{4(2\frac{e^2}{m^2} - 1)^2} \left\{ \frac{4}{\sin^4 \theta} - \frac{3}{\sin^2 \theta} + \frac{\left(\frac{e^2}{m^2} - 1\right)^2}{\left(2\frac{e^2}{m^2} - 1\right)^2} \left(1 + \frac{4}{\sin^2 \theta}\right) \right\} d\theta. \quad (36.5)$$

Let us write v and e in (36.5) in terms of the velocity v_1 and energy ϵ_1 of the incident electron in that system in which one of the electrons is at rest before the collision ($v_1 = 0$). From the invariance of the square of the total momentum p^2 of the system, it follows immediately that

$$\frac{e^{(c)}}{m} = \sqrt{\frac{1}{2} \left(\frac{\epsilon_1^{(0)}}{m} + 1 \right)}.$$

(The index c indicates the center-of-mass system, and the index 0 the system in which one of the electrons is at rest before the scattering.) Thus¹⁾

$$d\sigma = \frac{2\alpha^2}{m^2 v_1^4} \frac{\epsilon_1 + 1}{\left(\frac{\epsilon_1}{m}\right)^2} \left\{ \frac{4}{\sin^4 \theta} - \frac{3}{\sin^2 \theta} + \frac{\left(\frac{\epsilon_1}{m} - 1\right)^2}{4 \left(\frac{\epsilon_1}{m}\right)^2} \left(1 + \frac{4}{\sin^2 \theta}\right) \right\} d\theta, \quad (36.6)$$

where ϵ_1 is the initial electron energy in the "laboratory" system ($\epsilon_1 = m$) and θ is the scattering angle in the center-of-mass system. The latter is related to the scattering angle θ_0 in the laboratory system by the expression

$$\cos \theta = \frac{2 - \left(\frac{\epsilon_1}{m} + 1\right) \sin^2 \theta_0}{2 + \left(\frac{\epsilon_1}{m} - 1\right) \sin^2 \theta_0}.$$

The scattering angle in the center-of-mass system is uniquely related to the electron energy loss as a result of collision with the electron at rest. Let Δ be the ratio between the energy that the moving electron transfers to the electron at rest and the kinetic energy of the moving electron:

¹⁾ Möller, Ann. d. Phys. 14, 521 (1932). For the radiative corrections to Möller's formula, see A. Akhiezer and P. Polovin, Proc. Acad. Sci. USSR 90, 55 (1953).

$$\Delta = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 - m} = \frac{\epsilon'_2 - m}{\epsilon_1 - m}.$$

From conservation of energy and momentum it follows that

$$\Delta = \frac{1}{2} (1 - \cos \theta).$$

Inserting this expression into (36.6), we obtain

$$d\sigma = \frac{2\pi\alpha^2}{m v_1^4 (\epsilon_1 - m) \Delta^2 (1 - \Delta)^2} \left\{ 1 - \left[3 - \left(\frac{x-1}{x} \right)^2 \right] \Delta (1 - \Delta) + \left(\frac{x-1}{x} \right)^2 \Delta^2 (1 - \Delta)^2 \right\}, \quad x = \frac{\epsilon_1}{m}. \quad (36.7)$$

Equation (36.7) gives the energy distribution of the secondary electrons when a fast electron passes through matter. For small Δ ,

$$d\sigma \sim \frac{d\Delta}{\Delta^2}.$$

2. Positron-Electron Scattering.

Equation (36.1) can also be used for the collision of a positron with an electron. In this case, in correspondence with the previous section, we must perform the substitution

$$\begin{aligned} u'_1 &\rightarrow v_2, & p'_1 &\rightarrow -p_2^0, \\ u'_2 &\rightarrow v_1, & p'_2 &\rightarrow -p_1^0, \end{aligned}$$

where $\bar{p}_1^{(i)}$ and $\bar{p}_2^{(f)}$ are the initial and final four-momenta of the positron, and v_1 and v_2 are the amplitudes of the corresponding initial and final negative-frequency wave functions. Thus,

$$-M = \frac{(\bar{u}_2 \gamma_\mu u_1) (\bar{v}_1 \gamma_\mu v_2)}{(p_2 - p)^2} - \frac{(\bar{v}_1 \gamma_\mu u_1) (\bar{u}_2 \gamma_\mu v_2)}{(p_1^0 + p)^2} \quad (36.8)$$

(see Fig. 49).

is ⁹⁾ The calculation of $\Sigma |M|^2$ can be carried out in the same way as in the case of two electrons. The result

$$d\sigma = \left(\frac{\pi}{4\epsilon}\right)^2 \left\{ \left[\left(\frac{\epsilon}{m}\right)^2 - 1 \right]^2 \sin^4 \frac{\theta}{2} \left[1 + 4 \left(\frac{\epsilon^2}{m^2} - 1\right) \cos^2 \frac{\theta}{2} + 2 \left(\frac{\epsilon^2}{m^2} - 1\right)^2 \times \right. \right. \\ \left. \left. \times \left(1 + \cos^2 \frac{\theta}{2}\right) \right] + \left(\frac{m}{\epsilon}\right)^2 \left[3 + 4 \left(\frac{\epsilon^2}{m^2} - 1\right) + \left(\frac{\epsilon^2}{m^2} - 1\right)^2 (1 + \cos^2 \theta) \right] - \right. \\ \left. - \frac{m^4}{\epsilon^2 (\epsilon^2 - m^2) \sin^2 \frac{\theta}{2}} \left[3 + 4 \left(\frac{\epsilon^2}{m^2} - 1\right) (1 + \cos \theta) + \left(\frac{\epsilon^2}{m^2} - 1\right)^2 (1 + \cos \theta)^2 \right] \right\} d\theta,$$

where ϵ and θ are the energy and scattering angle in the center-of-mass system.

The energy distribution of the secondary electrons is given by

$$F(x; \Delta) = \frac{1}{x(x+1)} \left\{ \left[1 + 2(x-1)(1-\Delta) + (x-1)^2 \left(1 - x + \frac{x^2}{2}\right) \right] + \right. \\ \left. + \frac{(x-1)^2 \Delta^2}{(x+1)^2} \left[3 + 2(x-1) + (x-1)^2 \left(\frac{1}{2} - \Delta + \Delta^2\right) \right] - \right. \\ \left. - \frac{(x-1)\Delta}{x+1} \left[3 + 4(x-1)(1-\Delta) + (x-1)^2 (1-\Delta)^2 \right] \right\}, \quad (36.9)$$

where

$$x = \frac{e^{(p)}}{m}, \quad \Delta = \frac{e_A - m}{e^{(p)} - m}$$

(the electron is at rest before the collision).

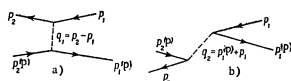


Fig. 49

⁹⁾ Hobbs, Proc. Roy. Soc. (London) 154, 195 (1936).

§ 37. Interaction Formula for Two Electrons up to Terms in $\frac{v^2}{c^2}$.

1. The Breit Formula.

In Section 35 we obtained an expression for the matrix element of the effective interaction energy for two electrons. This expression is of the form

$$U_{i \rightarrow f} = U_{AB; CD} = U_{AB; DC}, \quad (37.1)$$

where A and B are the initial, and C and D are the final, electron states. According to (35.15), $U_{AB; CD}$ can be written ¹⁰⁾

$$U_{AB; CD} = e^2 \int \psi_A^*(r_1) \psi_B^*(r_2) \frac{1 - \alpha_1 \alpha_2}{|r_1 - r_2|} e^{i(\omega_A r_1 - \omega_B r_2)} \psi_C(r_1) \psi_D(r_2) dr_1 dr_2. \quad (37.2)$$

Here α_1 acts on $\psi_A(r_1)$, and α_2 on $\psi_B(r_2)$. Let us examine the structure of expression (37.2). It contains an operator which depends on both the coordinates and the spins of both particles, namely $\frac{1 - \alpha_1 \alpha_2}{|r_1 - r_2|}$, and a "delay factor" $e^{i(\omega_A r_1 - \omega_B r_2)}$. This last factor which depends explicitly on the initial and final energies of the system, makes it impossible to introduce the general concept of the two-electron interaction operator U , whose matrix elements would be

$$U_{AB; CD} = \int \psi_A^*(r_1) \psi_B^*(r_2) U \psi_C(r_1) \psi_D(r_2) dr_1 dr_2. \quad (37.3)$$

For low velocities ($\frac{v}{c} \ll 1$, where c is the velocity of light), however, such an operator can be constructed. In order to do this we shall expand the matrix element in powers of $\frac{v}{c}$ up to terms in $\frac{v^2}{c^2}$. As in the similar expansion of Section 14, we shall not use the system of units in which $c = 1$. This has the advantage that the transition to the nonrelativistic approximation can be performed formally by an expansion in powers of $\frac{1}{c}$.

Up to terms in $\frac{1}{c^2}$, the expansion of the delay factor can be written

$$e^{i(\omega_A r_1 - \omega_B r_2)} = \frac{1}{|r_1 - r_2|} + i \frac{\omega_A - \omega_B}{c} |r_1 - r_2| - \frac{(\omega_A - \omega_B)^2}{2c^2} |r_1 - r_2|^2 + \dots \quad (37.4)$$

¹⁰⁾ Here and below the charge e is given in Gaussian (not Heaviside) units.

Since the matrix elements of α are of order of magnitude $\frac{v}{c}$, and therefore, $\alpha_1 \alpha_2 \sim \frac{v^2}{c^2}$, it is sufficient to keep only the first term of (37.4), where the delay factor enters in combination with $\alpha_1 \alpha_2$ in (37.2). The remaining terms are not multiplied by $1 - \alpha_1 \alpha_2$, but simply by unity. The second term in (37.4) vanishes after integration due to the orthogonality of ψ_A and ψ_B . The third term can be written symmetrically by making use of the fact that $\omega_A - \omega_C = \omega_D - \omega_B$, and therefore,

$$= (\omega_A - \omega_D)^2 |r_1 - r_2| + (\omega_A - \omega_D)(\omega_B - \omega_D) |r_1 - r_2|.$$

Further, the frequencies can be eliminated from the expression for the matrix element by making use of the Dirac equation

$$H(1)\psi_A(r_1) = m_1\psi_A(r_1)$$

and similar relations for the other wave functions. Bearing in mind that (37.4) is multiplied on the right by $\psi_A(r_1)\psi_B(r_2)$ and on the left by $\psi_C^*(r_1)\psi_D^*(r_2)$, and is integrated over dr_1 and dr_2 , we can replace ω_A and ω_B by the operators $H(1)$ and $H(2)$ on the right of the factor $|r_1 - r_2|$, and ω_C and ω_D can be replaced by $H(1)$ and $H(2)$ on the left of $|r_1 - r_2|$, that is $(\omega_A - \omega_C)(\omega_B - \omega_D) |r_1 - r_2| \rightarrow [H(1) - H(2)] |r_1 - r_2| H(1) H(2) + H(1) H(2) |r_1 - r_2| - H(1) |r_1 - r_2| H(2) - H(2) |r_1 - r_2| H(1) - [H(1), [H(2), |r_1 - r_2|]]$ (the square brackets denote commutators).

Since the Hamiltonian is

$$H = c\alpha p - \beta mc^2 + e\hat{A}\psi,$$

where $\hat{A}^{(e)}$ is the external field, the only term that does not commute with $|r_1 - r_2|$ is $\alpha \cdot p$. By calculating

$$c^2 [\alpha_1 p_1, |\alpha_2 p_2, |r_1 - r_2|],$$

we find that when (37.4) is inserted into (37.2), we can perform the substitution

$$\rightarrow \frac{(\omega_A - \omega_D)^2}{c^2} |r_1 - r_2| \rightarrow \frac{\alpha_1 \alpha_2}{c^2} \frac{(\alpha_1 p_1) (\alpha_2 p_2)}{|r_1 - r_2|}, \quad (37.5)$$

¹⁾ Since $\psi^\dagger \alpha \psi = \psi^\dagger \sigma_3 \psi + \chi^\dagger \sigma_3 \chi$, and $\chi \sim \frac{1}{c} \psi$ (see Section 14).

where

$$n = \frac{r_1 - r_2}{|r_1 - r_2|}.$$

Now inserting (37.5) and (37.4) into (37.2), we see that $U_{AB; CD}$ is actually the matrix element of the operator¹⁾

$$U = \frac{e^2}{|r_1 - r_2|} - \frac{e^2}{2} \frac{\alpha_1 \alpha_2 + (\alpha_1 n) (\alpha_2 n)}{|r_1 - r_2|}, \quad (37.6)$$

and that it satisfies (37.3).

The first term in (37.6) gives the Coulomb interaction. The second term gives the corrections due to motion of the charges.

Since this expression is correct only to terms of order $\frac{v^2}{c^2}$, the electron wave functions

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$$

may also be used only to that accuracy; this means that we may use relation (14.6), namely

$$\chi = \frac{cp}{2mc} \varphi.$$

Inserting this approximation for χ as well as (37.6) into (37.3), we obtain

$$\begin{aligned} U_{AB; CD} = e^2 \int & \varphi_C^*(r_1) \varphi_D^*(r_2) \varphi_A(r_1) \varphi_B(r_2) \frac{dr_1 dr_2}{|r_1 - r_2|} + \\ & + \frac{e^2}{4m^2 c^2} \int \{ [\sigma_1 p_1 \varphi_C(r_1)]^\dagger [\sigma_1 p_1 \varphi_A(r_1)] \varphi_D^*(r_2) \varphi_B(r_2) + \\ & + \varphi_C^*(r_1) \varphi_A(r_1) [\sigma_2 p_2 \varphi_D(r_2)]^\dagger [\sigma_2 p_2 \varphi_B(r_2)] \} \frac{dr_1 dr_2}{|r_1 - r_2|} + \\ & + \frac{e^2}{8m^2 c^2} \int \{ \varphi_C^*(r_1) \sigma_1 (\sigma_1 p_1) \varphi_A(r_1) + [\sigma_1 p_1 \varphi_C(r_1)]^\dagger \sigma_1 \varphi_A(r_1) \} \times \\ & \times \{ \varphi_D^*(r_2) \sigma_2 (\sigma_2 p_2) \varphi_B(r_2) + [\sigma_2 p_2 \varphi_D(r_2)]^\dagger \sigma_2 \varphi_B(r_2) \} \frac{dr_1 dr_2}{|r_1 - r_2|} - \\ & - \frac{e^2}{8m^2 c^2} \int \{ \varphi_C^*(r_1) (\sigma_1 n) (\sigma_1 p_1) \varphi_A(r_1) + \\ & + [\sigma_1 p_1 \varphi_C(r_1)]^\dagger (\sigma_1 n) \varphi_A(r_1) \} \{ \varphi_D^*(r_2) (\sigma_2 n) (\sigma_2 p_2) \varphi_B(r_2) + \\ & + [\sigma_2 p_2 \varphi_D(r_2)]^\dagger (\sigma_2 n) \varphi_B(r_2) \} \frac{dr_1 dr_2}{|r_1 - r_2|}. \end{aligned} \quad (37.7)$$

¹⁾ G. Breit, Phys. Rev. 34, 553 (1929); L. Landau, Physik. Z. Sowjetunion 8, 487 (1932); H. Bethe and E. Fermi, Z. Physik 77, 296 (1932).

We now introduce the approximate two-component electron wave function ψ , which is related to φ by [see (14.10), (14.13)]¹⁾

$$\varphi = \left(1 - \frac{p^2}{8mc^2}\right)\psi \quad (37.8)$$

and transform (37.7) so that it takes on a form similar to (37.3). In other words, we shall find an operator $U_{AB, CD}^E$, such that

$$U_{AB, CD} = \int \psi_C^*(r_1) \psi_D^*(r_2) U^E \psi_A(r_1) \psi_B(r_2) dr_1 dr_2. \quad (37.9)$$

To write (37.7) in the form of (37.9), we must integrate by parts. Since the integrand contains the factor $\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$, we must separate out the origin (the point $\mathbf{r}_1 = \mathbf{r}_2$). The integral over a surface surrounding the origin remains finite as $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow 0$. Since this integral depends only on the value of the integrand at $\mathbf{r}_1 = \mathbf{r}_2$, it can also be written in the form of a volume integral of an expression containing $\delta(\mathbf{r}_1 - \mathbf{r}_2)$. Finally, we obtain the following expression for the interaction energy operator for two electrons:

$$U^E = \frac{e^2}{r} - \pi \frac{e^2}{m^2 c^2} \delta(r) - \frac{e^2}{4m^2 c^2} \frac{1}{r^3} (|\mathbf{r} p_1| \sigma_1 - |\mathbf{r} p_2| \sigma_2 + 2 |\mathbf{r} p_1| \sigma_2 - 2 |\mathbf{r} p_2| \sigma_1) - \frac{e^2}{2m^2 c^2} \left(\frac{1}{r} p_1 p_2 + \frac{1}{r^3} \mathbf{r} (\mathbf{r} p_1) p_2 \right) + \frac{e^2}{4m^2 c^2} \left[\frac{\sigma_1 \sigma_2}{r^3} - \frac{3 (\sigma_1 \mathbf{r}) (\sigma_2 \mathbf{r})}{r^5} - \frac{8\pi}{3} \sigma_1 \sigma_2 \delta(r) \right], \quad r = r_1 = r_2. \quad (37.10)$$

The δ -function in this expression arises for the reason described above. We may then neglect the seeming divergence (since the neighborhood of $r = 0$ is already accounted for) of several integrals such as those which arise in inserting the last expression in the square brackets of (37.10) into (37.9). In order to avoid the divergences, one may first integrate over angles.

Naturally, the δ -function in (37.10) does not indicate a strong interaction. Those terms which contain the δ -function are multiplied by $\frac{1}{c^2}$, and thus they should be treated as small in comparison with the first term (the Coulomb interaction).

We shall discuss the interpretation of the separate factors of (37.10) in the next section in applying this equation to concrete problems. We note here that it bears some resemblance to expression (14.15) for the electron energy operator up to terms in $\frac{v^2}{c^2}$ in an external field. Thus, the second term in (37.10) is similar to the last

¹⁾ We note that (37.8) need be used only for the first term of (37.7), in the other terms, which already contain the factor $\frac{1}{c^2}$, it is sufficient to replace φ by ψ .

term in (14.15) with A_0 representing the Coulomb potential $A_0 = \frac{e}{r}$. This term is followed in (37.10) by terms of the same type as the next to last in (14.15) with $\mathbf{E}$ replaced by the Coulomb field $\mathbf{E} = \frac{e\mathbf{r}}{r^3}$. The last term in (37.10) corresponds to the last one in (14.7) where $\mathbf{H}$ is the field of a magnetic dipole whose moment is $\frac{e}{2mc} \boldsymbol{\sigma}$.

2. The Schroedinger Equation for a Two-Electron System.

The exchange interaction matrix element $U_{AB, CD}^E$ in (37.1) can be put into the form of (37.9) in exactly the same way as was $U_{AB, CD}$, so that

$$U_{AB, CD} = \int \psi_D^*(r_1) \psi_C^*(r_2) U^E \psi_A(r_1) \psi_B(r_2) dr_1 dr_2, \quad (37.11)$$

where U^E is the same operator (37.10).

If we use the antisymmetrized two-electron wave functions

$$\Phi_{AB}(r_1, r_2) = \frac{1}{\sqrt{2}} [\Psi_A(r_1) \Psi_B(r_2) - \Psi_B(r_1) \Psi_A(r_2)] \quad (37.12)$$

and similar functions for other pairs of states, then (37.1) can be written

$$U_{I \rightarrow I'} = \int \Phi_{CD}^* U^E \Phi_{AB} dr_1 dr_2. \quad (37.13)$$

These wave functions (if A and B are arbitrary stationary electron states in a given external field) are a complete set of antisymmetric eigenfunctions of the operator

$$H(1) + H(2)$$

(where H , as above, is the Hamiltonian for a single electron).

Due to the interaction between the electrons, the state whose wave function is Φ_{AB} is not stationary. If at $t = -\infty$ the electrons are in state Φ_{AB} , then at $t = +\infty$ they may be in some other state Φ_{CD} ; the probability for this transition, as we have seen, is proportional to

$$\left| \int \Phi_{CD}^* U^E \Phi_{AB} dr_1 dr_2 \right|^2.$$

We may wish to find the stationary states of a two-electron system, taking the interaction between the electrons into account (with an accuracy of $\frac{v^2}{c^2}$). Such states ϕ_n are superpositions of the states of (37.12) and are eigenfunctions of the operator

$$H = H(1) + H(2) + U^e. \quad (37.14)$$

Thus, we have found the Schrodinger equation for two electrons in configuration space, namely

$$H\phi_n = E_n\phi_n, \quad (37.15)$$

where E_n is the energy of the stationary state. In addition to (37.15), the wave functions ϕ_n must satisfy the antisymmetry condition.

3. Electron-Positron Interaction.

Let us now go on to a determination of the interaction energy between an electron and a positron up to terms of order $\frac{v^2}{c^2}$.

As was mentioned in Section 35 (see (35.7)), Equation (37.1) remains valid if the sign is changed in (37.2). Then if A is the initial, and C is the final, electron state as previously, D is now the initial, and B the final, positron state, and ψ_B and ψ_D are the corresponding negative-frequency wave functions. Instead of these wave functions, let us introduce the positron functions of (8.24), namely

$$\psi_B^{(p)} = C\psi_B^{(-)*}, \quad \psi_D^{(p)} = \psi_D^{(p)} C$$

and let us make use of the fact that

$$\bar{\psi}_D^{(-)} \gamma_4 \psi_B^{(-)} = \bar{\psi}_B^{(p)} \gamma_4 \psi_D^{(p)}.$$

Then the matrix element $U_{AB, CD}$ can be written

$$U_{AB, CD} = -e^2 \int \psi_B^{(-)*}(r_1) \psi_D^{(p)}(r_2) \frac{1 - \alpha_1 \alpha_2}{|r_1 - r_2|} e^{i \frac{\omega_A \omega_D}{c} |r_1 - r_2|} \times \psi_A(r_1) \psi_C^{(p)}(r_2) dr_1 dr_2. \quad (37.16)$$

In (37.16), as in (37.2), the final-state wave functions are on the left, and the initial-state wave functions are on the right; all these are positive-frequency wave functions. Therefore, in exactly the same way as (37.2) was transformed to (37.11), Equation (37.16) can be written in the form

$$U_{AB, CD} = - \int \psi_C^{(-)*}(r_1) \psi_D^{(p)*}(r_2) U^e \psi_A(r_1) \psi_B^{(p)}(r_2) dr_1 dr_2, \quad (37.17)$$

where U^e is the operator given by (37.10).

4. Exchange Interaction Between an Electron and Positron.

After expansion in powers of $\frac{1}{c^2}$, the exchange-matrix element takes on an entirely different form than for the case of two electrons:

$$U_{AB, D\bar{C}} = -e^2 \int \psi_D^{(-)*}(r_1) \psi_C^{(p)*}(r_2) \frac{1 - \alpha_1 \alpha_2}{|r_1 - r_2|} e^{\frac{i}{c} |\omega_A - \omega_D| |r_1 - r_2|} \times \psi_A(r_1) \psi_B^{(-)}(r_2) dr_1 dr_2. \quad (37.18)$$

In view of the fact that ω_D and ω_B are negative, the "delay factor" now contains not a difference, but a sum of frequencies

$$\omega_A - \omega_D = \varepsilon_A + \varepsilon_B^{(p)},$$

including the rest energy $2mc^2$. Thus, $\frac{\omega_A - \omega_D}{c}$ is not a small quantity of order $\frac{1}{c}$, but is a large quantity of order $\frac{1}{c}$, and therefore, the expansion (37.4) is no longer valid.

The transition to the low-velocity approximation is then easily obtained by replacing the energy in the exponent by the rest energy, according to

$$\varepsilon_A + \varepsilon_B^{(p)} \approx 2mc^2.$$

We then obtain

$$U_{AB, D\bar{C}} = -e^2 \int \psi_D^{(-)*}(r_1) \psi_C^{(p)*}(r_2) \frac{1 - \alpha_1 \alpha_2}{|r_1 - r_2|} e^{2imc |r_1 - r_2|} \psi_A(r_1) \psi_B^{(-)}(r_2) dr_1 dr_2. \quad (37.19)$$

We note that in (37.19) the initial and final states of the particles enter only through their wave functions, which means that this expression is no longer of the same form as (37.3).

The delay factor in (37.19) contains a large coefficient $\frac{2mc}{\hbar}$ in the exponent, and therefore, oscillates with a wave length $\frac{1}{2mc}$ (the "Compton wave length" is $\frac{1}{mc}$). The significant region of integration is $|\mathbf{r}_1 - \mathbf{r}_2| \sim \frac{1}{mc}$. Since the wave functions of the electron and positron (in agreement with the assumption of low velocities) do not vary significantly over this distance, we can separate out the integration over $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, assuming that in the arguments of the wave functions $\mathbf{r}_1 = \mathbf{r}_2$:

$$U_{AB; DO} = -e^2 \int \psi_D^{(+) *}(r_1) \psi_C^{(+)}(r_1) (1 - \alpha_1 \alpha_2) \psi_A(r_1) \psi_B^{(-)}(r_1) dr_1 \int \frac{e^{i2mc r}}{r} dr.$$

This last improper integral can be calculated by the usual method for calculating expressions containing oscillating integrands, namely

$$\int \frac{e^{i2mc r}}{r} dr = \lim_{\lambda \rightarrow 0} \int e^{i2mc r - \lambda r} \frac{dr}{r} = -\frac{\pi}{(mc)^2},$$

after which the exchange-matrix element can be written

$$U_{AB; DO} = \pi \left(\frac{e}{mc} \right)^2 \int \psi_D^{(+) *}(r_1) \psi_C^{(+)}(r_1) (1 - \alpha_1 \alpha_2) \psi_A(r_1) \psi_B^{(-)}(r_1) dr_1, \quad (37.20)$$

or

$$U_{AB; DO} = \pi \left(\frac{e}{mc} \right)^2 \int \psi_D^{(+) *}(r_1) \psi_C^{(+)}(r_2) (1 - \alpha_1 \alpha_2) \delta(r_1 - r_2) \psi_A(r_1) \psi_B^{(-)}(r_2) dr_1 dr_2.$$

Let us write the negative-frequency wave functions in (37.20) in terms of the positron wave functions. We then obtain

$$U_{AB; DO} = \frac{\pi e^2}{m^2 c^2} \int \{ (\psi_D^{(+) *} C \psi_A) (\psi_C^{(+)} C \psi_B^{(+)}) - (\psi_D^{(+) *} C \alpha \psi_{11}) (\psi_C^{(+)} C \alpha \psi_B^{(+)}) \} dr_1.$$

This last expression now contains the coefficient $\frac{1}{c^2}$. Therefore, we may drop χ in $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$, and in the desired approximation φ can be replaced by the two-component wave function Φ . Then all the four-by-four matrices

$$U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix},$$

where $\Gamma_1, \Gamma_2, \dots$ are two-by-two matrices, can be replaced by Γ_1 . Thus,

$$\left. \begin{aligned} C &\rightarrow 0, \\ \sigma_x C &\rightarrow \sigma_x, \\ \sigma_y C &\rightarrow i, \\ \sigma_z C &\rightarrow -\sigma_z. \end{aligned} \right\} \quad (37.21)$$

Making use of (37.21), we obtain

$$U_{AB; DO} = \pi \left(\frac{e}{mc} \right)^2 \int \{ \psi_D^{(+) *}(\lambda_1) \psi_C^{(+)}(\lambda_2) [1 - \sigma_{1z} \sigma_{2z} + 1 - \sigma_{1x} \sigma_{2x} + \psi_A(\lambda_1) \psi_B^{(+) *}(\lambda_2)] \} dr_1,$$

where λ_1 and λ_2 are spin variables (σ_1 acts on λ_1 , and σ_2 on λ_2).

In order to write the exchange-matrix element in a form similar to (37.17), let us transform the integrand of $\frac{U_{AB; DO}}{U_{AB; DC}}$ to the form

$$\Phi_A^{*}(\lambda_1) \Phi_B^{(+) *}(\lambda_2) T \Phi_A(\lambda_1) \Phi_B^{(+) }(\lambda_2).$$

It is easy to see that the operator T must be of the form

$$T = c_1 + c_2 \sigma_1 \sigma_2,$$

where c_1 and c_2 are constants. Indeed, the matrix element of T should be a scalar (in three-space). The operator T , which is a general bilinear combination of the four two-by-two matrices $(1, \sigma_x, \sigma_y, \sigma_z)$, satisfies this condition.

Making direct use of explicit forms for σ_x , we obtain

$$c_1 = \frac{3}{2}, \quad c_2 = \frac{1}{2}.$$

Thus,

$$U_{AB;DO} = \int \psi_0^*(r_1) \psi_B^{(0)}(r_2) U^{(ex)} \psi_A(r_1) \psi_D^{(0)}(r_2) dr_1 dr_2,$$

where⁹⁾

$$U^{(ex)} = \frac{e^2}{2} \left(\frac{e^2}{mc} \right)^2 (3 + \sigma_1 \sigma_2) \delta(r_1 - r_2). \quad (37.22)$$

The transition matrix element $U_{\underline{1} \rightarrow \underline{1}}$ can now be written

$$U_{\underline{1} \rightarrow \underline{1}} = \int \psi_0^* U^T \psi_{AD} dr_1 dr_2, \quad (37.23)$$

where

$$U^T = -U^e + U^{(ex)} \quad (37.24)$$

is the total interaction operator between an electron and a positron, and

$$\psi_{AD} = \psi_A(r_1) \psi_D^{(0)}(r_2), \quad \psi_{CB} = \psi_C(r_1) \psi_B^{(0)}(r_2). \quad (37.25)$$

The functions $\psi_{\underline{A}\underline{D}}$ ($\underline{A}$ and $\underline{D}$ are the quantum numbers of arbitrary electron and positron states in the given external field) are a complete set of eigenfunctions of the operator

$$H^{(e)}(1) + H^{(p)}(2),$$

where $H^{(e)}$ is the electron Hamiltonian, and $H^{(p)}$ is the positron Hamiltonian (in an external field they differ in the sign of the charge).

⁹⁾ Pirenne, Arch. sci. (Geneva) 29, 207 (1947), V. Berestetsky and L. Landau, J. Exptl.-Theoret. Phys. 19, 673 (1949).

The wave function $\Phi_{\underline{n}}(r_1, r_2)$ of a stationary electron-positron state satisfies the Schrodinger equation

$$[H^{(e)}(1) + H^{(p)}(2) + U^T] \Phi_{\underline{n}} = E_{\underline{n}} \Phi_{\underline{n}}, \quad (37.26)$$

and is not restricted by any symmetry requirements. The electron and positron in Equation (37.26) are treated as entirely distinct particles. The symmetry of the system with respect to changing the sign of the charge of both particles, however, leads to a definite "charge parity," as was shown in Section 19, paragraph 3 (see (19.25), (19.26)). In particular, symmetric states $[\Phi_{\underline{n}}(1, 2) = \Phi_{\underline{n}}(2, 1)]$; here 1 and 2 denote the coordinates and spins] have odd charge parity, and the antisymmetric states $[\Phi_{\underline{n}}(1, 2) = -\Phi_{\underline{n}}(2, 1)]$ have even charge parity.

§ 38. Positronium.

1. The Hamiltonian and Unperturbed Equation.

The Schrodinger equation (37.26) can be applied to the positronium problem, i.e., to bound states of the electron-positron system.

The interaction energy operator is given by (37.24), (27.22), and (37.10), and the electron and positron Hamiltonians are given by (14.15) with $\underline{A}_0 = 0$ (no external field).

The Hamiltonian of Equation (37.26) consists of terms of different orders of magnitude. We shall write it in the form

$$H = H^{(0)} + H^{(1)},$$

where $H^{(0)}$ consists of terms not containing $\frac{1}{c^2}$,

$$H^{(0)} = \frac{1}{2m} (p_1^2 + p_2^2) - \frac{e^2}{|r_1 - r_2|}, \quad (38.1)$$

and $H^{(1)}$ is proportional to $\frac{1}{c^2}$.

The problem of finding the positronium energy levels can be solved by perturbation theory. The operator $H^{(0)}$ may be considered the unperturbed Hamiltonian, and $H^{(1)}$ can then be treated as a perturbation.

The unperturbed problem is extremely simple, reducing to the hydrogen atom problem in nonrelativistic quantum mechanics. Let us use a system of coordinates in which the center of mass of positronium is at rest. Then

$$p_1 = -p_2 = p,$$

where $\mathbf{p}$ is the momentum operator corresponding to the relative position vector

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2.$$

The unperturbed Schroedinger equation is

$$\mathbf{p}^2 \Phi = m \left(E + \frac{e^2}{r} \right) \Phi. \quad (38.2)$$

This is the same as the equation for the motion of the electron in the hydrogen atom, where the mass is replaced by $\frac{m}{2}$, the reduced mass of two electrons. Thus, the energy levels of positronium lie half as high as the corresponding ones in hydrogen, and the orbit radii are twice as large.

The unperturbed levels, as is well known, depend only on the principal quantum number n , being independent of l and k , the total and orbital angular momenta.

2. Perturbation Operator.

Let us now go on to a consideration of the fine structure of positronium, that is to the splitting and shift of the energy levels due to the perturbation operator $H^{(1)} \Phi$. We shall write $H^{(1)}$ as the sum of five terms

$$H^{(1)} = V_1 + V_2 + V_3 + V_4 + V_5. \quad (38.3)$$

Here V_1 is the correction of order $1/c^2$ to the kinetic energy of the particles. According to (14.16)

$$V_1 = -\frac{1}{8m^2c^2} (\mathbf{p}_1^2 + \mathbf{p}_2^2) = -\frac{\mathbf{p}^4}{4m^2c^2}. \quad (38.4)$$

The remaining terms are in the interaction energy operator U^T . We shall lump all those terms of (37.34) which do not contain spin operators into V_2 . Then V_3 gives the interaction related to the orbital motion³

$$V_2 = \pi \left(\frac{e\hbar}{mc} \right)^2 \delta(r) + \frac{e^2}{2\pi^2 c^2} \left\{ \frac{1}{r} \mathbf{p}_1 \mathbf{p}_2 + \frac{1}{r^3} \mathbf{r} (\mathbf{r} \mathbf{p}_1) \mathbf{p}_2 \right\}.$$

³ V. Berestetsky, J. Exptl.-Theoret. Phys. 19, 1130 (1949).

⁴ We shall here use ordinary units, so that Planck's constant $\hbar$ appears explicitly; e is the charge in ordinary Gaussian units.

The operator V_3 can be written also in the following form, which exhibits its Hermiticity:

$$V_3 = \pi \left(\frac{e\hbar}{mc} \right)^2 \delta(r) - \frac{e^2}{2m^2c^2} \left\{ \frac{1}{r} \mathbf{p}^2 + \mathbf{p}^2 \frac{1}{r} - 4\pi\hbar^2 \delta(r) - \frac{1}{r^3} (\mathbf{r} \mathbf{p})^2 \right\}.$$

We shall write V_3 also in one other form, which is more convenient for future calculations:

$$V_3 = \pi \left(\frac{e\hbar}{mc} \right)^2 \delta(r) - \frac{i\hbar c^2}{m^2 c^2 r^3} \mathbf{r} \mathbf{p} - \frac{e^2}{m^2 c^2} \frac{1}{r} \mathbf{p}^2 + \frac{1}{2} \left(\frac{e\hbar}{mc} \right)^2 \frac{1}{r^3} L^2, \quad (38.5)$$

where

$$\hbar L = [\mathbf{r} \mathbf{p}]$$

is the orbital-angular-momentum operator.

We shall lump all terms which contain both momentum and spin operators into V_5 , so that

$$V_5 = \left(\frac{e\hbar}{2mc} \right)^2 \frac{1}{\hbar^2} \left\{ [\mathbf{r} \mathbf{p}_1] \sigma_1 - [\mathbf{r} \mathbf{p}_2] \sigma_2 + 2 [\mathbf{r} \mathbf{p}_1] \sigma_2 - 2 [\mathbf{r} \mathbf{p}_2] \sigma_1 \right\}.$$

This operator gives the spin-orbit coupling. It can be written

$$V_5 = \frac{3}{2} \left(\frac{e\hbar}{mc} \right)^2 \frac{1}{r^3} L s, \quad (38.6)$$

where

$$s = \frac{1}{2} (\sigma_1 + \sigma_2)$$

is the spin operator of the system.

The operator V_4 describes the interaction of the spin magnetic moments of the electron and positron, namely,

$$V_4 = \left(\frac{e\hbar}{2mc} \right)^2 \left[\frac{3}{r^3} (s_1 r) (s_2 r) - \frac{s_1 s_2}{r^3} \right] + \frac{8\pi}{3} \left(\frac{e\hbar}{2mc} \right)^2 s_1 s_2 \delta(r).$$

If we make use of the spin operators s , then V_4 can be written

$$V_4 = 6 \left(\frac{e\hbar}{2mc} \right)^2 \frac{1}{r^3} s_1 s_2 \left(\frac{x_1 x_2}{r^2} - \frac{1}{3} \delta_{12} \right) + \frac{8\pi}{3} \left(\frac{e\hbar}{2mc} \right)^2 (2s^2 - 3) \delta(r). \quad (38.7)$$

Finally, V_5 is the exchange interaction operator (37.22):

$$V_5 = U^{(ex)}.$$

It can also be written

$$V_5 = \pi \left(\frac{e\hbar}{mc} \right)^2 s^2 \delta(r). \quad (38.8)$$

It is important to note that $H^{(1)}$, as can be seen from (38.3) - (38.8), contains operators which act on the spin variables only in the form of the total spin s of the system. This means that the perturbation operator is diagonal with respect to the eigenvalues of s^2 , which are $s^2 = s(s+1)$. Therefore, the positronium levels can be divided into singlets and triplets, or parapositronium ($s=0$) and orthopositronium ($s=1$).

3. Fine Structure.

The operators V_1 and V_2 are diagonal with respect to the orbital quantum numbers l and are independent of the spin states. The interaction described by these operators removes the degeneracy in l . To find the corrections to the energy levels due to these operators, we need find only their expectation values ($\bar{V}_1$ and $\bar{V}_2$) in unperturbed states. Using the unperturbed wave equation (38.2), we easily arrive at

$$\bar{V}_1 = -\frac{1}{4mc^2} (E^2 + 2Ee^2 r^{-1} + e^4 r^{-2}) - \frac{e^2}{4\pi^2 r^3} \left(p^2 \frac{1}{r} - \frac{1}{r} p^2 \right).$$

The last term vanishes. In fact,¹⁾

¹⁾ Actually

$$\frac{1}{R} \frac{1}{r^3} \langle r p \rangle = \int \Phi^* \frac{1}{r^3} \frac{\partial \Phi}{\partial r} dr = \int_0^\infty R \frac{dR}{dr} dr = -\frac{1}{2} R^2(0),$$

where R is the radial part of the wave function; since $R(0)$ is nonzero only when $l=0$, we have $\bar{R}^2(0) = 4\pi \Phi^2(0)$.

$$p^2 \frac{1}{r} - \frac{1}{r} p^2 = 4\pi \hbar^2 \delta(r) + \frac{2i\hbar}{r^3} r p, \\ \bar{\delta}(r) = -\frac{i}{2\pi \hbar} \frac{1}{r^3} \langle r p \rangle = \Phi^2(0). \quad (38.9)$$

We can calculate the various expectation values from the known results for hydrogen (taking into account the reduced mass):

$$\left. \begin{aligned} E &= -\frac{e^2}{4a_0^2}, \quad \Phi^2(0) = \frac{1}{8\pi a_0^3 n^3}, \\ \bar{r}^{-1} &= \frac{1}{2a_0 n^2}, \quad \bar{r}^{-2} = \frac{1}{2a_0^2 n^3 (2l+1)}, \\ \bar{r}^{-3} &= \frac{1}{4a_0^3 n^4 (l+1)(2l+1)}, \end{aligned} \right\} \quad (38.10)$$

where $a_0 = \frac{\hbar^2}{mc^2}$, a_0 is the hydrogen-atom radius, and n is the principal quantum number.

Using these values, we obtain

$$\bar{V}_1 = -\alpha^2 e_0 \frac{1}{8\pi^2} \left(\frac{1}{2l+1} - \frac{3}{8\pi} \right), \quad \alpha = \frac{e^2}{\hbar c} = \frac{1}{137}. \quad (38.11)$$

Similarly, according to (38.5) and making use of (38.2) and (38.9), we obtain²⁾

$$\bar{V}_2 = 3\pi \left(\frac{e\hbar}{mc} \right)^2 \Phi^2(0) - \frac{e^2}{mc^2} \left(\frac{E + \frac{e^2}{r}}{r} \right) + \frac{1}{2} \left(\frac{e\hbar}{mc} \right)^2 l(l+1) \bar{r}^{-3}.$$

Inserting (38.10) into this expression, we obtain

²⁾ Here $\frac{1}{r} \frac{\partial}{\partial r}$ must be considered zero when $l=0$, in spite of the fact that as $l \rightarrow 0$, $\frac{1}{r} \frac{\partial}{\partial r}$ diverges when $l=0$. Indeed, in deriving the expression for $\bar{U}^0$, those terms in whose matrix elements the origin was significant were considered separately. In the product $l \frac{1}{r} \frac{\partial}{\partial r}$, the factor l comes from integration over angles, and this, as was noted in Section 37, may be performed first in the singular terms.

$$\bar{V}_2 = a^2 e_0 \left[\frac{1}{4n^4} \delta_{ll} - \frac{1}{8a^3} \left(\frac{3}{2l+1} - \frac{1}{n} \right) \right]. \quad (38.12)$$

The operator V_3 is also diagonal with respect to l and s . It vanishes in the singlet state, and depends on the total angular momentum j in the triplet states. Making use of the fact that

$$2ls = j(j+1) - l(l+1) - s(s+1),$$

we can write immediately

$$\bar{V}_3 = \begin{cases} \frac{3}{16} a^2 e_0 \frac{j(j+1) - l(l+1) - 2}{n^4(l+1)(2l+1)} & \text{for } s=1, \quad l \neq 0, \\ 0 & \text{for } s=0 \text{ or } l=0. \end{cases} \quad (38.13)$$

We shall treat the two terms of expression (38.7) for V_4 separately. The second of these is diagonal with respect to s and is nonzero only when $l=0$. It determines the expectation value of V_4 when $l=0$, namely

$$\bar{V}_4 = \frac{8\pi}{3} \left(\frac{e\hbar}{mc} \right)^2 \frac{1}{4} [2s(s+1) - 3] \psi^2(0), \quad (38.14)$$

$$\bar{V}_4 = \frac{1}{4n^3} a^2 e_0 \begin{cases} -1 & \text{for } s=0, \\ \frac{1}{3} & \text{for } s=1 \quad (l=0). \end{cases}$$

The first term in (38.7) has matrix elements which are not diagonal with respect to l . From the conservation laws for the total angular momentum, spin angular momentum, and parity, it is easy to see that the only nonzero nondiagonal elements of V_4 are those connecting states with equal values of j and M for $s=1$ and $l=j \pm 1$.

To calculate the matrix elements of the first term of (38.7), let us write it in the form

$$6 \left(\frac{e\hbar}{2mc} \right)^2 u(r) \Lambda(\sigma_1 \sigma_2),$$

where

$$u = \frac{1}{r^3} (n_p n_q - \frac{1}{3} (-1)^p \delta_{pq} - q),$$

$$\Lambda = (-1)^{p+q} s_{-p} s_{-q}$$

and $\underline{n}$ is a unit vector; we shall denote its components by [see (3.23), (3.24)]

$$n_0 = n_z;$$

$$n_{\pm 1} = \pm \frac{1}{\sqrt{2}} (n_x \pm i n_y);$$

the components of $\underline{s}$ are denoted similarly. The product $\underline{u} \Lambda$ denotes summation over p and q from -1 to $+1$.

Let us expand the wave function $\psi_{j, l, M}$ for a given state characterized by the quantum numbers j, l, s , $\underline{M}$ in products of orbital and spin functions $\psi_{l, m}$ and $\chi_{s, \mu}$:

$$\psi_{j, l, M} = \sum_{m, \mu} C_{lm; sp}^{jM} \psi_{lm} \chi_{sp}.$$

The desired matrix element, which we shall denote by $(u\Lambda)_{l' m' l m}^j$, can be written

$$(u\Lambda)_{l' m' l m}^j = \sum_{m_1, \mu_1} \sum_{m_2, \mu_2} C_{lm_1; sp}^{jM} C_{l'm_2; sp}^{jM} (u)_{l' m' l m}^{jM} (\Lambda)_{sp, sp}^{jM},$$

where

$$(u)_{l' m' l m}^{jM} = \int \psi_{l' m'}^* u \psi_{l m} d\tau,$$

$$\Lambda_{sp, sp}^{jM} = \chi_{sp}^* \Lambda \chi_{sp}.$$

In order to calculate $(u)_{l' m' l m}^j$, we note that its angular dependence is simply related to the spherical function Y_{lm} , namely,

¹⁾ The coefficients C are the same as in (4.4).
²⁾ See Section 55.

$$n_p n_q = \frac{1}{3} (-1)^p \delta_{p, -q} = a_{pq} V_2, \quad p+q=0,$$

where

$$a_{0,0} = \sqrt{\frac{4\pi(2+\eta)(2-\eta)}{3 \cdot 3 \cdot 5}}, \quad a_{\pm 1,0} = \sqrt{\frac{4\pi(2 \pm q)(3 \pm q)}{2 \cdot 3 \cdot 5}}.$$

The orbital wave functions ψ and ψ' contain spherical functions Y_{lm} and Y'_{lm} ($|l-l'| = 0, 2$). We shall integrate the matrix element $(\psi | V | \psi')$ over angles by using the expansion

$$Y_{2r} Y_{lm} = \sum_{\lambda, \mu} c_{\lambda\mu}^{(2)} Y_{l+\lambda, m+\mu},$$

where

$$c_0 = \sqrt{\frac{5l(l+1)}{4\pi(2l-1)(2l+3)}}, \quad c_1 = \sqrt{\frac{15l(l+1)(l+2)}{8\pi(2l+3)(2l+5)}}, \\ c_{-1} = \sqrt{\frac{15l(l-1)}{8\pi(2l-1)(2l-3)}}.$$

Then

$$(u)_{l'm'}^{lm} = (r^{-3})_{l'm'}^{lm} c_{l'm}^{(l'-l)} a_{pq},$$

where

$$(r^{-3})_{l'm'}^{lm} = \int \psi_{lm}^* \frac{1}{r^3} \psi_{l'm} dr.$$

In order to calculate the $(A)_{\mu}^{\mu}$ (we have suppressed the common index s) we shall use the matrix multiplication rule

$$\Lambda_{\mu}^{\mu} = (-1)^{p+q} \sum_{\nu} (s_{-p})_{\mu}^{\nu} (s_{-q})_{\nu}^{\mu},$$

where the $(s_{-p})_{\mu}^{\nu}$ are known angular momentum matrix elements. They differ from zero if

$$p = \mu - \nu, \quad q = \nu - \mu'.$$

Performing the necessary operations and inserting the values of the coefficients, we obtain the following expression for the diagonal elements of $\bar{V}_4$:

$$\bar{V}_4 = \frac{a^2 e_0}{8l(l+1)(2l+1)n^3} \begin{cases} 1 & \text{for } j=l, \\ -\frac{l}{2l+3} & \text{for } j=l+1, \quad (l \neq 0, s=1) \\ -\frac{l+1}{2l-1} & \text{for } j=l-1 \end{cases} \quad (38.15)$$

(when $l=0$, the only possibility is $j=l+1$, and thus, (38.15) vanishes since $\frac{l}{2l+3}$ is to be considered finite). The nondiagonal elements of V_4 are

$$(V_4)_{j+1}^{j-1} = a^2 e_0 \frac{\sqrt{j(j+1)}}{2j+1} (r^{-3})_{j+1}^{j-1}. \quad (38.16)$$

We shall not calculate $(r^{-3})_{j+1}^{j-1}$, since in general the expressions obtained are not simple, and in all special cases involving the discrete spectrum, their calculation is elementary. In particular, when $j=1$ and $n=3$, expression (38.16) vanishes, so that for the first three principal quantum numbers the orbital angular momentum is a good quantum number.

For singlet states, as is easily shown from (38.7), when $l \neq 0$

$$\bar{V}_4 = 0.$$

Finally, $\bar{V}_4$ is nonzero only for $l=0, s=1$:

$$\bar{V}_4 = \begin{cases} 0 & \text{for } s=0, \\ \frac{a^2}{4n^3} e_0 \delta_{l_0} & \text{for } s=1. \end{cases} \quad (38.17)$$

Thus, of all the terms contained in the perturbation operator $H^{(1)}$, only V_4 has nondiagonal elements, which refer to

$$l=j+1, \quad l'=j-1, \quad s=1.$$

Therefore, states with $l = \frac{1}{2}$ [with parity $(-1)^{l+1}$], both singlet and triplet, can be classified according to their orbital angular momentum. Let us denote the shift of each such level, referred to its unperturbed value, by

$w(\frac{2j+1}{2}, \frac{1}{2})$ (the spectroscopic designation of the term is given in parentheses). Then

$$w(\frac{1}{2}, \frac{1}{2}) = -a^2 \frac{1}{2n^3} \left(\frac{1}{2j+1} - \frac{11}{32n} \right), \quad (38.18)$$

$$w(\frac{3}{2}, \frac{1}{2}) = w(\frac{1}{2}, \frac{1}{2}) - \frac{a^2}{8n^3(j+1)(2j+1)}. \quad (38.19)$$

Terms with parity $(-1)^{\frac{1}{2}+1}$ are superpositions of unperturbed terms for which $l = \frac{1}{2} \pm 1$. For these terms

$$w = \frac{\bar{V}_+ - \bar{V}_-}{2} \pm \sqrt{\left(\frac{\bar{V}_+ - \bar{V}_-}{2} \right)^2 + |(V_{j+1}^{j-1})|^2}, \quad (38.20)$$

where

$$V_- = \bar{V}(j-1)_j = w((j-1)_{j-1}) + a^2 \frac{2}{3n^2} \frac{1}{2j-1} \left(\frac{3}{2} - \frac{1}{2j+1} \right),$$

$$V_+ = \bar{V}(j+1)_j = w((j+1)_{j+1}) - \frac{a^2}{4n^2(j+1)(2j+3)} \left(\frac{3}{2} + \frac{1}{2j+1} \right)$$

and $\bar{V}_\pm = \frac{1}{2} \bar{V}_\pm$ is defined by (38.16).

These formulas make it possible, in particular, to determine the energy differences between the ground states of ortho- and parapositronium. The difference in the s -dependence of the energy for $l = 0$ is contained only in $\bar{V}_+$ as given by (38.14) and $\bar{V}_-$ as given by (38.17). It follows from these equations that the ground level of orthopositronium is above that of parapositronium, and that the energy difference between them is

$$\Delta = \frac{1}{4} a_0 a^2 \left(\frac{4}{3} + 1 \right) = 8.2 \cdot 10^{-4} \text{ eV}. \quad (38.21)$$

(The first term in the parentheses is due to the magnetic interaction V_μ , and the second is due to the exchange term V_g).

4. Zeeman Effect.

Let us examine the Zeeman effect in positronium, which is quite unique.¹⁾ The orbital magnetic moment of positronium is always zero. Indeed, the appropriate term in the expression for the interaction with a magnetic field is

$$\frac{e}{2mc} H [r_1 p_1] - \frac{e}{2mc} H [r_2 p_2] = 0,$$

since $[r_1 p_1] = [r_2 p_2]$. Therefore, the perturbation energy is diagonal with respect to l . Further, since spin magnetic moment

$$\mu = \frac{e\hbar}{2mc} (\sigma_1 - \sigma_2)$$

is not proportional to the spin angular momentum of the system

$$s = \frac{1}{2} (\sigma_1 + \sigma_2),$$

the eigenvalues of these two operators are in general different. Of the four eigenfunctions $\chi_{\pm\mu}$ which belong to given eigenvalues $\pm$ and μ of the spin and its projection,

$$\chi_{11} = \alpha(1)\alpha(2),$$

$$\chi_{1,-1} = \beta(1)\beta(2),$$

$$\chi_{10} = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) + \beta(1)\alpha(2)],$$

$$\chi_{00} = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)],$$

[where α and β are spin functions of a single particle, the number (1) refers to an electron, and the number (2) refers to a positron], the first two χ_{11} and $\chi_{1,-1}$ are simultaneous eigenfunctions of μ_z belonging to the eigenvalue $\mu_z = 0$. These states do not interact with the magnetic field. The functions χ_{10} and χ_{00} are not eigenfunctions of μ_z ; the only combinations of these which are such eigenfunctions are

$$\chi = c_1 \chi_{10} + c_0 \chi_{00}. \quad (38.22)$$

¹⁾ V. Berestetsky, J. Exptl.-Theoret. Phys. 19, 1130 (1949).

In other words, the matrix μ_z with the basis $\chi_{\pm\mu}$ can be written

$$\mu_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{e\hbar}{mc} \\ 0 & 0 & \frac{e\hbar}{mc} & 0 \end{pmatrix}$$

Thus, the matrix for the energy or interaction with the field contains only such elements as connect singlet states with triplet states. But the perturbation energy V , as we have seen, actually does not have such elements.

Therefore, in weak fields (that is, when $\frac{e\hbar}{mc} H \ll W^{(2)} \dots - W^{(4)} \dots$) there exists no effect proportional to the field. In the other limiting case of an extremely strong field, the superposition (38.22) takes place, and the splitting is $\pm \frac{e\hbar}{mc} H$.

It is simple to write out the equation for the splitting in a magnetic field for the ground state (as well as for the s -states with $n = 2$ and 3). The wave functions of the stationary states are the superpositions

$$\chi^{(i)} = c_1^{(i)} \chi_{10} + c_0^{(i)} \chi_{00}, \quad i = 1, 2 \quad (38.23)$$

($i = 1$ corresponds to a pure triplet state in the absence of the magnetic field, and $i = 2$ corresponds to a singlet state). The coefficients $c_{\pm}^{(i)}$ satisfy the equations

$$\left. \begin{aligned} (E^{(i)} - E_{10}) c_1^{(i)} + \frac{e\hbar}{mc} H c_0^{(i)} &= 0, \\ \frac{e\hbar}{mc} H c_1^{(i)} + (E^{(i)} - E_{00}) c_0^{(i)} &= 0, \end{aligned} \right\} \quad (38.23')$$

where E_{10} and E_{00} are the energies of the triplet and singlet states, respectively, in the absence of a magnetic field, and $E^{(i)}$ is the energy in the field H .

The condition that (38.23') be solvable can be written

$$E^{(i)} = \frac{E_{10} + E_{00}}{2} \pm \sqrt{\left(\frac{E_{10} - E_{00}}{2}\right)^2 + \left(\frac{e\hbar}{mc} H\right)^2}. \quad (38.24)$$

$\left| \frac{c_1^{(i)}}{c_0^{(i)}} \right|^2$ gives the relative amounts of triplet and singlet states in a given stationary state in the magnetic field.

The magnetic field influences the decay probability of positronium. Let λ_0 and λ_1 be the decay probabilities for para- and orthopositronium. In a magnetic field the decay probability of positronium in the state $\chi^{(1)}$ is

$$\lambda^{(1)} = |c_1^{(1)}|^2 \lambda_1 + |c_0^{(1)}|^2 \lambda_0, \quad (|c_1|^2 + |c_0|^2 = 1). \quad (38.25)$$

Even a small value of the coefficient c_0 will greatly increase the decay probability of the state $\chi^{(1)}$ (triplet state with no magnetic field), since (see Section 33)

$$\lambda_0 \gg \lambda_1.$$

In a weak magnetic field

$$\left| \frac{c_1^{(1)}}{c_0^{(1)}} \right|^2 = \frac{\left(\frac{e\hbar}{mc} H \right)^2}{(E_{10} - E_{00})^2}. \quad (38.26)$$

Measurements of the effect of the magnetic field on positronium decay made possible the experimental determination of the energy difference $\Delta = E_{10} - E_{00}$ of the ground states of ortho- and parapositronium, using Equation (38.26)⁹. The measured value of Δ agrees with the theoretical one given by (38.21).

§ 39. Internal Conversion of Gamma Rays.

1. Expansion of Retarded Potentials in Spherical Waves.

Due to its electromagnetic interaction with electrons, a nucleus in an excited state can undergo transition to a lower energy state by transferring its excitation energy to the electrons in the atomic shells or by creating an electron-positron pair. This process is called internal conversion of gamma rays (this terminology indicates that the excited nucleus may also lose energy by emission of photons). We can treat internal conversion as a special case of the interaction of a nuclear proton with an electron, as described in Section 35, using the general expression (35.15).

Let us denote the initial state of the electron by ψ_1 , and its final state by ψ_2 ; similarly for the proton we shall use φ_1 and φ_2 . According to (35.8) and (35.15) the matrix element of the effective perturbation energy is

$$\begin{aligned} U_{1 \rightarrow 2} &= -\alpha \int \bar{\psi}_2(r_1) \gamma_1 \psi_1(r_1) \frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|} \bar{\varphi}_2(r_2) \gamma_2 \varphi_1(r_2) dr_1 dr_2 = \\ &= -\alpha \int (\bar{\psi}_2(r_1) \gamma_2 \psi_1(r_1)) (\bar{\varphi}_2(r_2) \gamma_1 \varphi_1(r_2)) \frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|} dr_1 dr_2, \end{aligned} \quad (39.1)$$

⁹ M. Deutch and E. Duff, Phys. Rev. 84, 601 (1951).

where

$$\omega = E_1 - E_2 = \epsilon_2 - \epsilon_1,$$

$E_{1,2}$ are the proton energies and $\epsilon_{1,2}$ the electron energies.

The calculation of these integrals is simplified due to the difference in the sizes of the regions in which the proton and electron motions take place. For the proton this region is of the order of the nuclear radius. For the electron, on the other hand, the region is much larger. Therefore, in integrating (39.1) we may assume that the contribution comes primarily from the region in which $r_1 < r_2$.

When this condition is satisfied, we may write¹⁾

$$\frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|} = \frac{i\omega}{4\pi} \sum_{lm} \varphi_{lm}(r_1) \Phi_{lm}(r_2), \quad (39.2)$$

where

$$\left. \begin{aligned} \varphi_{lm} &= g_l(\omega r) Y_{lm}\left(\frac{r}{r_1}\right), \\ \Phi_{lm} &= G_l(\omega r) Y_{lm}\left(\frac{r}{r_2}\right), \\ g_l(z) &= (2\pi)^{1/2} i^l \frac{J_{l+1/2}(z)}{\sqrt{z}}, \\ G_l(z) &= (2\pi)^{1/2} i^l \frac{H_{l+1/2}^{(1)}(z)}{\sqrt{z}}, \end{aligned} \right\} \quad (39.3)$$

and J is a Bessel function, and $H^{(1)}$ is the Hankel function of the first kind. We shall show that we may also write

$$\alpha_1 \alpha_2 \frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|} = \frac{i\omega}{4\pi} \sum_{jlm} (\alpha_1 A_{jlm}^*(r_1)) (\alpha_2 B_{jlm}(r_2)), \quad (39.4)$$

¹⁾ Equation (39.2) is an addition theorem for cylindrical functions applied to the function

$$H_{1/2}^{(1)}(\omega|r_1-r_2|) \approx i \sqrt{\frac{2}{\pi}} \frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|}.$$

where

$$\left. \begin{aligned} A_{jlm} &= g_l(\omega r) Y_{lm}, \\ B_{jlm} &= G_l(\omega r) Y_{lm} \end{aligned} \right\} \quad (39.5)$$

[$\mathbf{Y}_{jlm}$ is a spherical vector, defined in (4.11)]. Indeed, every vector function can be expanded in a series of spherical vectors $\mathbf{Y}_{jlm}$. Let α_1 be an arbitrary vector independent of $\mathbf{r}_2$. Then

$$\frac{\alpha_1 e^{i\omega|r_1-r_2|}}{|r_1-r_2|},$$

treated as a function of $\mathbf{r}_2/r_2$, can be written

$$\alpha_1 \frac{e^{i\omega|r_1-r_2|}}{|r_1-r_2|} = \frac{i\omega}{4\pi} \sum_{jlm} C_{jlm}^M(r_1, r_2) Y_{jlm}\left(\frac{\mathbf{r}_2}{r_2}\right). \quad (39.6)$$

In order to find the coefficients C_{jlm}^M , let us multiply this equation by $\mathbf{Y}_{j'lm}^* \cdot \mathbf{M} \left(\frac{\mathbf{r}_2}{r_2}\right)$ and integrate over all directions of the unit vector $\frac{\mathbf{r}_2}{r_2}$. Making use of (39.2) and the orthogonality property of the spherical vectors, as well as their relation to the spherical functions (see Section 4), we obtain

$$C_{jlm}^M = (\alpha_1 A_{jlm}^*(r_1)) G_l(\omega r_2).$$

Inserting this expression into (39.6) and multiplying by α_2 , we arrive at (39.4).

Instead of the sets $\mathbf{A}_{jlm}$ and $\mathbf{B}_{jlm}$ (for given values of j and m , the other index satisfies $l = j \pm 1, j$), we may use the set $\mathbf{A}_{jlm}^{(\lambda)}$, as defined by (30.6) and (30.16), and a similarly defined set $\mathbf{B}_{jlm}^{(\lambda)}$ (the difference between these is that G_l is replaced by G_l). The vectors $\mathbf{A}_{jlm}^{(\lambda)}$ and $\mathbf{B}_{jlm}^{(\lambda)}$, as was shown in Section 4, have the property that asymptotically (when $\omega r \gg 1$) they are transverse (for $\lambda = 0, 1$) or longitudinal (for $\lambda = -1$). It is simple to verify that

$$\sum_{l=j-1}^{j+1} (\alpha_1 A_{jlm}^*) (\alpha_2 B_{jlm}) = \sum_{l=-1}^1 (\alpha_1 A_{jlm}^{(0)}) (\alpha_2 B_{jlm}^{(0)}).$$

Inserting the above series into (39.1), we obtain

$$U_{i \rightarrow f} = -e^2 \sum_{jM} \left\{ (\psi_{jM})_{21} (\phi_{jM})_{21} - \sum_{\lambda=-1}^1 (\alpha A_{jM}^{(1)\lambda})_{21} (\alpha B_{jM}^{(1)\lambda})_{21} \right\}. \quad (39.7)$$

The parentheses denote those matrix elements of the corresponding quantities which connect electron or proton states as follows:

$$\begin{aligned} (\psi_{jM})_{21} &= \int \Psi_2^* \psi_{jM} \Psi_1 dr, \\ (\phi_{jM})_{21} &= \int \phi_2^* \phi_{jM} \phi_1 dr, \\ (\alpha A_{jM}^{(1)\lambda})_{21} &= \int \Psi_2^* \alpha A_{jM}^{(1)\lambda} \Psi_1 dr, \\ (\alpha B_{jM}^{(1)\lambda})_{21} &= \int \phi_2^* \alpha B_{jM}^{(1)\lambda} \phi_1 dr \end{aligned}$$

(In what follows we shall sometimes drop the indices on the parentheses).

Equation (39.7) can be simplified by using some of the relations obtained in Section 30. According to (30.7) and (30.11),

$$(\alpha A_{jM}^{(1)\lambda})_{21} = (\psi_{jM}^{\lambda})_{21}.$$

Furthermore, when the condition

$$\omega R \ll 1,$$

is satisfied, as we shall assume it to be (R is the nuclear radius),

$$A_{jM}^{(-1)} = -\sqrt{\frac{j}{j+1}} A_{jM}^{(0)}.$$

In view of this, three of the four terms in the braces of (39.7) can be combined as follows:

$$(\psi_{jM}^{\lambda})_{21} (\phi_{jM})_{21} - (\alpha A_{jM}^{(1)\lambda})_{21} (\alpha B_{jM}^{(1)\lambda})_{21} - (\alpha A_{jM}^{(0)})_{21} (\alpha B_{jM}^{(0)})_{21} = (\alpha A_{jM}^{(0)})_{21} (\alpha B_{jM}^{(0)})_{21}, \quad (39.8)$$

where

$$\begin{aligned} B_{jM}^{(0)} &= -\alpha B_{jM}^{(1)} - \sqrt{\frac{j}{j+1}} (\phi_{jM} - \alpha B_{jM}^{(1)}) = \\ &= -\sqrt{\frac{j}{j+1}} \phi_{jM} - \sqrt{\frac{2j+1}{j+1}} \alpha B_{j, j-1, M}. \end{aligned} \quad (39.9)$$

We shall henceforth also make use of the notation

$$B_{jM}^{(0)} = -\alpha B_{jM}^{(1)}. \quad (39.10)$$

We note that in spite of the fact that $B_{jM}^{(-1)}$ and ϕ_{jM} are related by an expression similar to (30.7), namely,

$$B_{jM}^{(-1)} = \frac{1}{\omega} \nabla \phi_{jM},$$

the matrix elements $(\phi_{jM})_{21}$ and $(\alpha B_{jM}^{(-1)})_{21}$ cannot be expressed in terms of each other in a way similar to (30.11), since a transformation of the kind made in Section 30 cannot be made in this case due to the singularity of ϕ_{jM} . [Because there is a pole at $r = 0$ when the integration by parts necessary on going from (30.10) to (30.11) is performed, there remains an integral over a surface enclosing the pole; this surface integral does not vanish as $r \rightarrow 0$. In other words, the operator H is not self-conjugate with respect to the functions ϕ_{jM} .]

Thus, (39.7) can be written

$$U_{i \rightarrow f} = -e^2 \sum_{jM} \{ (\alpha A_{jM}^{(1)\lambda})_{21} (B_{jM}^{(1)})_{21} + (\alpha A_{jM}^{(0)})_{21} (B_{jM}^{(0)})_{21} \}. \quad (39.11)$$

2. Conversion Coefficient.

Equation (39.11) for the internal conversion matrix element has a simple physical interpretation. Up to a constant coefficient, $(\alpha A_{jM}^{(1)\lambda})_{21}$ and $(\alpha A_{jM}^{(0)})_{21}$ are the matrix elements for emission of a photon with angular momentum j and parity $(-1)^j$ or $(-1)^{j+1}$; they are proportional to the matrix element of the electric or magnetic 2^j -pole moment of the nucleus (see Section 30). The operators $B_{jM}^{(\lambda)}$, whose electron-state matrix elements are taken, can be interpreted as the perturbation energy due to the field related to nuclear transitions.

The quantities $\Phi_{JM}^{(\lambda)}$ and $\Phi_{JM}^{(\lambda)}$ can be called the potentials of the electric or magnetic multipole. Indeed, the potentials $\Phi_{JM}^{(\lambda)}$ and $\Phi_{JM}^{(\lambda)}$, as can easily be shown, are solutions of the wave equation (35.17) in the form of outgoing spherical waves with a singularity at the origin. This singularity corresponds to the emission of an electric or magnetic multipole of J -th order (2^L -pole) at $r = 0$. The process of conversion consists, so to speak, of the absorption of these waves by the electrons in the vicinity of the nucleus.

We note that the expression for $\Phi_{JM}^{(\lambda)}$ differs from that for the matrix element of the operator

$$A_{JM} = \alpha A_{JM},$$

corresponding to absorption of a photon in an eigenstate of the angular momentum and parity, only in that the regular function G_L is replaced by G_L , which has a singularity at $r = 0$. We make note of the fact that the electric multipole potentials (39.9) have the structure of (5.20), corresponding to a definite choice of the constant C in the general expression (5.19) for the photon potentials. The arbitrary choice of C for A_{JM} , which is permitted in view of gauge invariance, is forbidden in this case due to the singularity of the potentials.

Let the angular momentum of the nucleus in the initial and final states be J_1 and J_2 . Then the matrix element $(\alpha A_{JM}^{(\lambda)})_{21}$ is nonzero when $|J_1 - J_2| \leq J \leq J_1 + J_2$. If $\omega R \ll 1$, only the lowest possible value of J is significant. For a given value of J the matrix elements $(\alpha A_{JM}^{(\lambda)})_{21}$ and $(\alpha A_{JM}^{(\lambda)})_{21}$ cannot both be nonzero, due to the parity selection rule. It is possible in principle for electric and magnetic multipoles whose orders differ by unity, for instance $(\alpha A_{JM}^{(1)})_{21}$ and $(\alpha A_{JM}^{(0)})_{21}$, both to participate in a process. We shall not, however, consider this possibility, since an evaluation of the matrix elements (see Section 30) leads to the conclusion that they are in general of different orders of magnitude. As a rule, therefore, internal conversion is determined by just one term of the sum (39.11). Thus, the general expression

$$w = 2\pi |U_{i \rightarrow f}|^2 \delta(E_1 - E_2 - \epsilon_1 - \epsilon_2)$$

for the conversion probability may be simplified.

When the angular momentum of the nucleus changes by ¹⁾

$$|J_1 - J_2| = L,$$

¹⁾ If $J_1 = J_2$ (and $J_1 \neq 0$), then the dipole transition ($L = 1$) is most probable. For the case $J_1 = J_2 = 0$ see Section 41.

and the parity changes by $(-1)^L$, the conversion probability is determined by the potential of the electric 2^L -pole, and is

$$w_L^{(1)} = \frac{\alpha^2}{8\pi} \omega^2 |(\alpha A_{LM}^{(1)})_{21}|^2 |(B_{LM}^{(1)})_{21}|^2 \delta(E_1 - E_2 - \epsilon_1 - \epsilon_2).$$

When the nuclear angular momentum changes by L and the parity changes by $(-1)^{L+1}$, the conversion probability is determined by the magnetic 2^L -pole and is

$$w_L^{(0)} = \frac{\alpha^2}{8\pi} \omega^2 |(\alpha A_{LM}^{(0)})_{21}|^2 |(B_{LM}^{(0)})_{21}|^2 \delta(E_1 - E_2 - \epsilon_1 - \epsilon_2).$$

The conversion coefficient is the ratio between the probability for conversion and the probability for radiation of a photon in a given nuclear transition. According to Section 30, the radiation probability is

$$w_{L\text{rad}}^{(1)} = \frac{\omega^4}{2\pi} |(\alpha A_{LM}^{(1)})_{21}|^2.$$

Thus, the conversion coefficient for an electric ($\lambda = 1$) or magnetic ($\lambda = 0$) 2^L -pole is

$$\beta_L^{(1)} = \frac{w_L^{(1)\text{conv}}}{w_{L\text{rad}}^{(1)}} = \frac{\alpha}{4} |(B_{LM}^{(1)})_{21}|^2 \delta(E_1 - E_2 - \epsilon_1 - \epsilon_2). \quad (39.12)$$

In Equation (39.12) it is assumed that the electron wave functions are normalized so that

$$\int \psi_1^* \psi_1 d\tau = \int \psi_2^* \psi_2 d\tau = 1.$$

If the wave function ψ_2 (belonging to the continuous spectrum) is normalized for a unit energy interval, so that

$$\int \psi_2^* \psi_2 d\tau = \delta(\epsilon - \epsilon'), \quad \text{whereas } \psi_1 \text{ remains normalized as above, then the expression for the conversion}$$

coefficient becomes

$$\beta_L^{(1)} = \frac{\alpha^2}{4} \left| \int \psi_2^* B_{LM}^{(1)} \psi_1 d\tau \right|^2. \quad (39.13)$$

3. K-shell; Reduction to Radial Integrals.

It is significant that the conversion coefficient $\beta_L^{(0)}$ does not contain quantities which depend on the properties of the nucleus. Its calculation necessitates only the knowledge of the electron wave functions. If the electron is initially in the K-shell, then the conversion coefficient is

$$\beta_L^{(0)} = \frac{a\omega}{4} \sum_K \sum_f |(B_{LM}^{(0)})_{fK}|^2.$$

Here λ is either 1 or 0, depending on whether the transition is an electric or magnetic one; Σ denotes summation over the two electrons of the K-shell, and Σ_K denotes summation over the possible final states allowed by the selection rules. The calculation may be simplified by making use of the fact that for a closed shell the conversion coefficient cannot depend on the quantum number M . Indeed, the latter depends on the choice of the z axis, and a closed shell is spherically symmetric. We may, therefore, find $\beta_L^{(\lambda)}$ on the assumption that $M = 0$. Further, since the perturbation is independent of the azimuth angle, when $M = 0$ the conversion probability must be the same for two electrons which differ only in the value of the quantum number $m = \pm \frac{1}{2}$. Therefore,

$$\beta_L^{(0)} = \frac{a\omega}{2} \sum_f |(B_{L0}^{(0)})_f|^2.$$

The final state must have an energy

$$e_f = \omega - |e_K|,$$

where $|e_K| = -\epsilon_K$ is the binding energy of the electron in the K-shell (internal conversion may take place if $\omega > |e_K|$).

The angular momenta of the final states can be found by using the conservation laws for angular momentum and parity. These laws show that the change in the angular momentum of the nucleus will be exactly the same as that when emitting a photon with angular momentum L and parity $(-1)^L$ or $(-1)^{L+1}$, and that the selection rules for the final electron state will be the same as those for absorption of such an electron. This can be shown directly also by considering the angular part of the integral (39.1) for the matrix element, which is the same as the angular part of the matrix element for photon absorption.

Since in the initial state

$$j_1 = 1/2, \quad l_1 = 0,$$

the final state for the case of an electric multipole satisfies

$$l_2 = L, \quad j_2 = l_2 \pm 1/2 = L \pm 1/2.$$

For a magnetic multipole

$$l_2 = L + 1, \quad j_2 = l_2 - \frac{1}{2} = L + 1/2,$$

or

$$l_2 = L - 1, \quad j_2 = l_2 + 1/2 = L - 1/2.$$

Thus,

$$\left. \begin{aligned} \beta_L^{(0)} &= \frac{a}{2} \omega \{ |(B_{L0}^{(0)})_{L+1/2, L}|^2 + |(B_{L0}^{(0)})_{L-1/2, L}|^2 \}, \\ \beta_L^{(0)} &= \frac{a}{2} \omega \{ |(B_{L0}^{(0)})_{L+1/2, L+1}|^2 + |(B_{L0}^{(0)})_{L-1/2, L-1}|^2 \}. \end{aligned} \right\} \quad (39.14)$$

(The two indices of the matrix elements denote the quantum numbers of the final electron state; the first gives j_2 , and the second l_2 .)

We shall use the general expression (11.6) [see also (11.9)] for the wave functions of the electron:

$$\psi_{jlm} = \begin{pmatrix} i g_s(r) \Omega_{jlm}(n) \\ f_s(r) \Omega_{j'l'm}(n) \end{pmatrix}, \quad (39.15)$$

where

$$n = \frac{r}{\rho}, \quad l' = 2j - l, \quad \Omega_{j'l'm} = (2n) \Omega_{jlm}.$$

In particular, the wave function of an electron in the K-shell ($j = 1/2, l = 0, m = -1$) can be written

$$\psi_K = \begin{pmatrix} ig_K(r) u \\ f_K(r) (\sigma n) u \end{pmatrix}, \quad (39.16)$$

where u is a constant spinor normalized according to

$$u^* u = \frac{1}{4\pi}.$$

Using the definition (39.9) of $B_{LM}^{(1)}$, we obtain

$$(B_{LM}^{(1)})_{j,j_s} = -\sqrt{\frac{L}{L+1}} (Q_{LM})_{j,j_s} - \sqrt{\frac{2L+1}{L+1}} (\alpha B_{LM, L-1, 0})_{j,j_s}. \quad (39.17)$$

Inserting expressions (39.15) and (39.16) for the wave functions, as well as (39.3) and (39.5) for the potentials, we obtain

$$(Q_{LM})_{j,j_s} = \int (g_s g_K + f_s f_K) G_L r^2 dr u_{j,j_s}^* u, \quad (39.18)$$

where we have introduced the notation

$$u_{j,j_s}^* = \int \Omega_{j,m}^* Y_{LM} d\sigma. \quad (39.19)$$

(We note that if $l = 0$ and $j = \frac{1}{2}$, then $u_{\frac{1}{2}, \frac{1}{2}}$ is the same as u .)

The second term in (39.17) contains the quantity

$$(\alpha B_{LM, L-1, 0})_{j,j_s} = -i \int g_s g_K Q_{L-1} r^2 dr \left(\int \Omega_{j,m}^* (\sigma n) \sigma \cdot Y_{L, L-1, 0} d\sigma \right) u + \\ + i \int g_s f_K Q_{L-1} r^2 dr \left(\int \Omega_{j,m}^* \sigma (\sigma n) \cdot Y_{L, L-1, 0} d\sigma \right) u. \quad (39.20)$$

The angle integral in this expression can be transformed to the form of (39.19). To do this let us resolve the spherical vector $\underline{\sigma}$, $\underline{L}$, $\underline{L}-1$, 0 into its transverse and longitudinal components

$$Y_{L, L-1, 0} = \sqrt{\frac{L}{2L+1}} Y_{L0}^{(-1)} - \sqrt{\frac{L+1}{2L+1}} Y_{L0}^{(1)}.$$

Since

$$(\sigma n) \sigma = n + i [\sigma n],$$

we may write

$$(\sigma n) \sigma Y_{L, L-1, 0} = -\sqrt{\frac{L}{2L+1}} n Y_{L0}^{(-1)} + \sqrt{\frac{L+1}{2L+1}} i \sigma [n Y_{L0}^{(1)}].$$

Writing the longitudinal and transverse spherical vectors in terms of spherical functions

$$Y_{L0}^{(-1)} = n Y_{L0}, \quad i [n Y_{L0}^{(1)}] = Y_{L0}^{(0)} = \frac{1}{\sqrt{L(L+1)}} L Y_{L0},$$

($L = -\frac{1}{2} [\mathbf{r} \nabla]$ is the orbital angular momentum operator), we obtain

$$(\sigma n) \sigma Y_{L, L-1, 0} = \left(-\sqrt{\frac{L}{2L+1}} + \frac{\sigma L}{\sqrt{L(L+1)}} \right) Y_{L0}.$$

Similarly

$$\sigma (\sigma n) Y_{L, L-1, 0} = \left(-\sqrt{\frac{L}{2L+1}} - \frac{\sigma L}{\sqrt{L(L+1)}} \right) Y_{L0}.$$

Further, making use of the fact that $L \sigma$ is self-conjugate, we can transform the integral

$$\int \Omega_{j,m}^* L \sigma Y_{L0} d\sigma$$

into a form where $\mathbf{L} \cdot \boldsymbol{\sigma}$ acts on $\Omega_{j_2 m_2}$, which is an eigenfunction of this operator with eigenvalue

$$j_2(j_2+1) - l_2(l_2+1) - \frac{3}{4}.$$

In this way we arrive at the following values for the angle integrals in (39.20):

$$\int \Omega_{j_2 m_2}^* (\sigma n) \sigma_{L-1,0} u d\sigma = \frac{u_{j_2}^* u}{\sqrt{L(2L+1)}} \begin{cases} 0 & \text{for } j_2 = L + \frac{1}{2}, \\ 2L+1 & \text{for } j_2 = L - \frac{1}{2}, \end{cases} \quad (39.21)$$

$$\int \Omega_{j_2 m_2}^* (\sigma n) \sigma_{L-1,0} u d\sigma = \frac{u_{j_2}^* u}{\sqrt{L(2L+1)}} \begin{cases} -2L & \text{for } j_2 = L + \frac{1}{2}, \\ 1 & \text{for } j_2 = L - \frac{1}{2}, \end{cases}$$

Inserting (39.21) into (39.20), and (39.18) into (39.17), we finally obtain the following expression for the matrix element which determines the internal conversion coefficient for an electric multipole:

$$\begin{aligned} (B_{L0}^{(1)})_{L+1/2, L} &= u_{L+1/2, L}^* \sqrt{\frac{L}{L+1}} (|R_1| + |R_2| + 2|R_4|), \\ (B_{L0}^{(1)})_{L-1/2, L} &= u_{L-1/2, L}^* \sqrt{\frac{L}{L+1}} \left(|R_1| + |R_2| - \frac{2L+1}{L} |R_3| - \frac{1}{L} |R_4| \right), \end{aligned} \quad (39.22)$$

$j_2 = L + \frac{1}{2},$
 $j_2 = L - \frac{1}{2},$

where the R_i denote the following radial integrals:

$$\begin{aligned} R_1 &= \int g_K G_L r^2 dr, \\ R_2 &= \int f_K G_L r^2 dr, \\ R_3 &= \int f_K G_{L-1} r^2 dr, \\ R_4 &= \int g_K G_{L-1} r^2 dr \end{aligned} \quad (39.23)$$

(the functions f_K and g_K refer to the final electron states).

The calculation of the quantities $\frac{u^*}{\sqrt{4\pi}} \frac{u}{0}$ which enter into (39.22) is easily performed by expressing the spherical spinor in (39.19) in terms of spherical functions according to (10.6). The quantum number m_1 (the projection of the electron angular momentum in the initial state) can be chosen as $\frac{1}{2}$ for the K -shell, as has already been mentioned. We then also have $m_2 = \frac{1}{2}$. Then

$$u = \frac{1}{\sqrt{4\pi}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$u_{j_2}^* u = \frac{1}{\sqrt{4\pi(2L+1)}} \begin{cases} \sqrt{L+1} & \text{for } j_2 = L + \frac{1}{2}, \\ \sqrt{L} & \text{for } j_2 = L - \frac{1}{2}. \end{cases} \quad (39.24)$$

Inserting (39.22) and (39.23) into (39.13), we obtain the following expression for the internal conversion coefficient of an electric multipole on the K -shell:

$$\begin{aligned} \beta_L^{(1)} &= \frac{\alpha\omega}{8\pi} \left\{ \frac{L}{2L+1} (|R_1| + |R_3| + 2|R_4|)_{j_2=L+1/2, l_2=L}^2 + \right. \\ &\quad \left. + \frac{L^2}{(2L+1)(L+1)} \left(|R_1| + |R_2| - \frac{2L+1}{L} |R_3| - \frac{1}{L} |R_4| \right)_{j_2=L-1/2, l_2=L}^2 \right\} \end{aligned} \quad (39.25)$$

The matrix element which gives the conversion of a magnetic multipole can be written

$$\begin{aligned} (B_{L0}^{(2)})_{j_2 l_2} &= -i \int f_K G_L r^2 dr \int \Omega_{j_2 m_2}^* (\sigma n) \sigma V_{L2}^{(0)} u d\sigma + \\ &\quad + i \int g_K G_L r^2 dr \int \Omega_{j_2 m_2}^* (\sigma n) \sigma V_{L2}^{(0)} u d\sigma. \end{aligned} \quad (39.26)$$

The angle integrals in this expression can also be reduced to integrals of the form of (39.19). We do this by writing $Y_{L0}^{(0)}$ in terms of the spherical function Y_{L0} and noting that

$$\Omega_{j_2 m_2}^* (\sigma n) = \Omega_{j_2' m_2'}^* \quad (j_2' = 2j - l_2)$$

so that

$$\int \Omega_{j_2 m_2}^*(\sigma n) \sigma Y_{j_2 m_2}^{(0)} d\sigma =$$

$$= \int \Omega_{j_2 m_2}^* \frac{L}{\sqrt{L(L+1)}} Y_{j_2 m_2} d\sigma = \frac{j_2(j_2+1) - l_2'(l_2'+1) - \frac{3}{4}}{L(L+1)} u_{j_2 l_2'}^* u.$$

The second integral in (39.26) differs from the first only in its sign.
Introducing the notation

$$\left. \begin{aligned} R_6 &= \int f_K G_L r^2 dr, \\ R_6 &= \int f_K G_L r^2 dr, \end{aligned} \right\} \quad (39.27)$$

we can write (39.26) in the form

$$(B_{j_2 l_2'}^{(0)})_{j_2 l_2'} = u_{j_2 l_2'}^* u (|R_6| + |R_6|) \frac{1}{\sqrt{L(L+1)}} \begin{cases} L & \text{for } j_2 = L - \frac{1}{2}, \\ -(L+1) & \text{for } j_2 = L - \frac{1}{2}, \end{cases}$$

where $\frac{u}{L+1} \frac{u}{L+1}$ is given by (39.24).

The internal conversion coefficient for a magnetic multipole on the K-shell is given by

$$\beta_K^{(0)} = \frac{\alpha \omega}{8\pi} \frac{1}{2L+1} \left\{ L(|R_6| + |R_6|)_{j_2=L+\frac{1}{2}, l_2=L+1}^2 + \right. \\ \left. + (L+1)(|R_6| + |R_6|)_{j_2=L-\frac{1}{2}, l_2=L-1}^2 \right\}. \quad (39.28)$$

4. K-Shell; Results.

The radial integrals (39.23) and (39.27) which enter into expressions (39.25) and (39.28) for the conversion coefficients can be calculated simply only in the limiting case of small nuclear charge $Z\alpha$ and high transition frequencies ω . In this case we may neglect both the binding energy of the electron in the K-shell and the effect of the nuclear charge on the continuous-spectrum electron wave function (that is, on the final state). In this approximation

$$\left. \begin{aligned} g_K &= 2(Z\alpha n)^{1/2}, \quad f_K = 0, \\ g_s &= \sqrt{\frac{\epsilon+m}{2}} \frac{J_{L+1/2}(pr)}{\sqrt{pr}}, \\ f_s &= \sqrt{\frac{\epsilon-m}{2}} \frac{J_{L+1/2}(pr)}{\sqrt{pr}}, \\ l_2' &= 2l_2 - l_2', \quad p = \sqrt{\epsilon^2 - m^2}, \quad \epsilon = \omega + m. \end{aligned} \right\} \quad (39.29)$$

When we insert (39.29) into (39.23) and (39.27), we obtain the following expressions for the radial integrals:

$$\left. \begin{aligned} R_1 &= (2\pi Z\alpha n)^{1/2} \sqrt{\frac{2(\epsilon+m)}{\omega}} I_L, \\ R_2 &= 0, \\ R_3 &= (2\pi Z\alpha n)^{1/2} \sqrt{\frac{2(\epsilon-m)}{\omega}} I_{L-1}, \\ R_4 &= 0, \\ R_6 &= (2\pi Z\alpha n)^{1/2} \sqrt{\frac{2(\epsilon-m)}{\omega}} I_L, \end{aligned} \right\}$$

where

$$I_L = \int J_{L+1/2}(pr) H_{L+1/2}^{(1)}(pr) r dr.$$

This integral can be evaluated as

$$I_L = \frac{2}{\pi i} \left(\frac{p}{\omega} \right)^{L+1/2} \frac{1}{p^2 - \omega^2}.$$

The conversion coefficients then become

$$\left. \begin{aligned} \beta_{j_2 l_2'}^{(0)} &= 2\gamma (Z\alpha)^2 \frac{(L+1)\omega^2 + 4Ln^2}{(L+1)\omega^2} \left(1 + \frac{2m}{\omega} \right)^{L-1/2}, \\ \beta_{j_2 l_2'}^{(0)} &= 2\gamma (Z\alpha)^2 \left(1 + \frac{2m}{\omega} \right)^{L+1/2}, \\ \left(\gamma \approx \frac{1}{137} \right). \end{aligned} \right\} \quad (39.30)$$

In the nonrelativistic approximation ($\omega \ll m$) these expressions become

$$\left. \begin{aligned} \beta_L^{(0)} &= \alpha (Z\alpha)^3 \frac{L}{L+1} \left(\frac{2m}{\omega} \right)^{L+1/2} \\ \beta_L^{(0)} &= \alpha (Z\alpha)^3 \left(\frac{2m}{\omega} \right)^{L+1/2} \end{aligned} \right\} \quad (39.31)$$

Equations (39.30) and (39.31) make it possible to achieve a qualitative understanding of the way in which the conversion coefficient depends on the energy and angular momentum. The $\beta_L^{(0)}$ are large for small ω and large L . Quantitatively, however, the free-electron approximation gives very poor accuracy even for small Z and large ω . The expressions for the β_L can be integrated if we use the nonrelativistic approximations for the electron wave functions.¹⁾ The accuracy of these results, however, is not great, since even for low energies the electron wave functions differ noticeably from their nonrelativistic approximations in the region of small r , and this region is most significant in the integrals for the β_L .

If the exact wave functions are used, the integration can only be performed numerically. The Table below presents the results of such numerical calculations²⁾ using Dirac wave functions for an electron in a Coulomb field.

5. Effect of the Finite Size of the Nucleus.

The calculations we have discussed do not take into account the finite size of the nucleus, which introduces two types of corrections to the matrix element for the conversion probability. In the first place, (39.2) is not valid for the region $r_2 < R$ (where R is the radius of the nucleus). Therefore, the exact expression for the matrix element involving transitions with a given value of L (that is, conversion of an electric or magnetic multipole) will not be of the form (39.11), but

$$U_{i \rightarrow f} = -\frac{e\omega}{4\pi} (aa^0)_{fi} \int_{r_2 > R} \psi_f^0(r_2) \psi_i^0(r_1) dr_2 - \int_{r_1 < R, r_2 < R} \bar{\psi}_f(r_1) \bar{\psi}_i(r_2) \gamma_{11} \gamma_{22} \psi_f(r_1) \psi_i(r_2) \frac{e^{i\omega(r_1-r_2)}}{|r_1-r_2|} dr_1 dr_2. \quad (39.32)$$

(In the expressions obtained previously, it was assumed that $R \rightarrow 0$.) In the second place, in the neighborhood of the nucleus the electron wave functions may differ significantly from the Coulomb wave functions, as we have seen in Section 12. As is shown by direct analysis, the main error arises from extending the integral in the first term of (39.32) into the region $0 < r_2 < R$, using the Coulomb wave functions in this region. The latter have a singularity at $r = 0$, and, therefore, for large Z , the region $0 < r_2 < R$ gives a contribution which is comparable to that from the region $R < r_2 < \infty$. The true wave functions, however, have no singularity at $r = 0$ (since $r \cdot \psi \rightarrow 0$ as $r \rightarrow 0$), and therefore, the integral over the region $r_2 < R$ should be small compared with that over the region $r_2 > R$. In practice, a good approximation is³⁾

K-SHELL INTERNAL CONVERSION COEFFICIENTS

(The number in parentheses gives the power of 10 which multiplies the figure given in the Table.)

Z	ω/m	$\beta_1^{(0)}$	$\beta_2^{(0)}$	$\beta_3^{(0)}$	$\beta_4^{(0)}$	$\beta_5^{(0)}$
10	0.3	9.041 (-4)	8.549 (-3)	6.822 (-2)	5.173 (-1)	3.845 (0)
	0.5	1.394 (-4)	1.112 (-3)	5.866 (-3)	2.947 (-2)	1.451 (-1)
	1.0	2.476 (-5)	8.837 (-5)	2.803 (-4)	8.544 (-4)	2.558 (-3)
	1.8	6.361 (-6)	1.521 (-5)	3.300 (-5)	7.221 (-5)	1.331 (-4)
	3.0	2.552 (-6)	4.746 (-6)	8.051 (-6)	1.368 (-5)	2.301 (-5)
	5.0	1.224 (-6)	1.820 (-6)	2.645 (-6)	3.786 (-6)	5.377 (-6)
20	0.3	5.617 (-3)	4.892 (-2)	3.576 (-1)	2.485 (0)	1.700 (1)
	0.5	1.183 (-3)	6.850 (-3)	3.302 (-2)	1.589 (-1)	7.395 (-1)
	1.0	1.721 (-4)	5.941 (-4)	1.817 (-3)	5.340 (-3)	1.541 (-2)
	1.8	4.522 (-5)	1.084 (-4)	2.371 (-4)	5.020 (-4)	1.047 (-3)
	3.0	1.808 (-5)	3.391 (-5)	5.544 (-5)	1.013 (-4)	1.701 (-4)
	5.0	8.542 (-6)	1.327 (-5)	1.872 (-5)	2.370 (-5)	4.116 (-5)
30	0.3	1.518 (-2)	1.194 (-1)	7.852 (-1)	4.919 (0)	3.042 (1)
	0.5	3.362 (-3)	1.804 (-2)	8.285 (-2)	3.627 (-1)	1.592 (0)
	1.0	5.175 (-4)	1.717 (-3)	5.014 (-3)	1.424 (-2)	3.955 (-2)
	1.8	1.393 (-4)	3.343 (-4)	7.233 (-4)	1.510 (-3)	3.100 (-3)
	3.0	5.552 (-5)	1.075 (-4)	1.907 (-4)	3.263 (-4)	5.476 (-4)
	5.0	2.583 (-5)	4.198 (-5)	6.417 (-5)	9.488 (-5)	1.375 (-4)
40	0.3	2.949 (-2)	2.042 (-1)	1.180 (0)	6.518 (0)	3.569 (1)
	0.5	6.850 (-3)	3.372 (-2)	1.410 (-1)	5.672 (-1)	2.255 (0)
	1.0	1.116 (-3)	3.538 (-3)	9.954 (-3)	2.698 (-2)	7.211 (-2)
	1.8	3.083 (-4)	7.407 (-4)	1.589 (-3)	3.275 (-3)	6.832 (-3)
	3.0	1.227 (-4)	2.457 (-4)	4.456 (-4)	7.613 (-4)	1.278 (-3)
	5.0	5.621 (-5)	9.603 (-5)	1.513 (-4)	2.277 (-4)	3.332 (-4)
54	0.3	5.640 (-2)	3.082 (-1)	1.414 (0)	6.252 (0)	2.776 (1)
	0.5	1.401 (-2)	5.830 (-2)	2.123 (-1)	7.487 (-1)	2.643 (0)
	1.0	2.479 (-3)	7.323 (-3)	1.943 (-2)	5.001 (-2)	1.277 (-1)
	1.8	7.136 (-4)	1.725 (-3)	3.682 (-3)	7.508 (-3)	1.590 (-2)
	3.0	2.848 (-4)	6.026 (-4)	1.115 (-3)	1.948 (-3)	3.539 (-3)
	5.0	1.277 (-4)	2.366 (-4)	3.910 (-4)	6.040 (-4)	8.964 (-4)
64	0.3	7.931 (-2)	3.456 (-1)	1.283 (0)	4.663 (0)	1.735 (1)
	0.5	2.072 (-2)	7.485 (-2)	2.435 (-1)	7.808 (-1)	2.547 (0)
	1.0	3.910 (-3)	1.099 (-2)	2.824 (-2)	7.075 (-2)	1.766 (-1)
	1.8	1.165 (-3)	2.856 (-3)	6.135 (-3)	1.248 (-2)	2.475 (-2)
	3.0	4.676 (-4)	1.041 (-3)	1.974 (-3)	3.464 (-3)	5.834 (-3)
	5.0	2.009 (-4)	4.113 (-4)	7.067 (-4)	1.112 (-3)	1.666 (-3)
72	0.3	9.882 (-2)	3.425 (-1)	1.037 (0)	3.143 (0)	9.969 (0)
	0.5	2.699 (-2)	8.579 (-2)	2.555 (-1)	7.651 (-1)	2.361 (0)
	1.0	5.394 (-3)	1.469 (-2)	3.718 (-2)	9.185 (-2)	2.258 (-1)
	1.8	1.656 (-3)	4.161 (-3)	9.014 (-3)	1.841 (-2)	3.628 (-2)
	3.0	6.703 (-4)	1.572 (-3)	3.654 (-3)	5.490 (-3)	9.087 (-3)
	5.0	2.941 (-4)	6.255 (-4)	1.112 (-3)	1.776 (-3)	2.673 (-3)
78	0.3	1.136 (-1)	3.196 (-1)	8.102 (-1)	2.095 (0)	5.732 (0)
	0.5	3.215 (-2)	9.242 (-2)	2.594 (-1)	7.417 (-1)	2.190 (0)
	1.0	6.712 (-3)	1.810 (-2)	4.575 (-2)	1.121 (-1)	2.716 (-1)
	1.8	2.120 (-3)	5.499 (-3)	1.211 (-2)	2.463 (-2)	4.819 (-2)
	3.0	8.659 (-4)	2.135 (-3)	4.233 (-3)	7.516 (-3)	1.261 (-2)
	5.0	3.783 (-4)	8.549 (-4)	1.562 (-3)	2.518 (-3)	3.795 (-3)

¹⁾ For the corresponding results and literature referring to it see Nuclear Spectroscopy by L. Groshov and L. Shapiro, (State Tech. Press, 1952).

²⁾ M. Rose, G. Goertzel, B. Spinard, J. Hiser and P. Strong, Phys. Rev. 83, 79 (1951).

³⁾ L. Sliv, J. Exptl.-Theoret. Phys. 21, 770 (1951).

(Continued)

Z	ω/m	$\xi_1^{(0)}$	$\xi_2^{(0)}$	$\xi_3^{(0)}$	$\xi_4^{(0)}$	$\xi_5^{(0)}$
83	0.5	3.669 (-2)	9.721 (-2)	2.615 (-1)	7.1F4 (-1)	2.022 (0)
	1.0	7.980 (-3)	2.163 (-2)	5.483 (-2)	1.372 (-1)	3.171 (-1)
	1.8	2.584 (-3)	6.970 (-3)	1.554 (-2)	3.153 (-2)	6.109 (-2)
	3.0	1.063 (-3)	2.708 (-3)	5.584 (-3)	9.931 (-3)	1.659 (-2)
	5.0	4.646 (-4)	1.115 (-3)	2.083 (-3)	3.379 (-3)	5.095 (-3)
88	0.5	4.138 (-2)	1.020 (-1)	2.643 (-1)	6.900 (-1)	1.813 (0)
	1.0	9.409 (-3)	2.013 (-2)	6.691 (-2)	1.594 (-1)	3.098 (-1)
	1.8	3.132 (-3)	8.929 (-3)	2.016 (-2)	4.663 (-2)	7.762 (-2)
	3.0	1.306 (-3)	3.625 (-3)	7.432 (-3)	1.321 (-2)	2.189 (-2)
	5.0	5.097 (-4)	1.468 (-3)	2.805 (-3)	4.568 (-3)	6.868 (-3)
92	0.5	4.516 (-2)	1.066 (-1)	2.675 (-1)	6.590 (-1)	1.601 (0)
	1.0	1.068 (-2)	3.083 (-2)	7.878 (-2)	1.850 (-1)	4.172 (-1)
	1.8	3.040 (-3)	1.102 (-2)	2.507 (-2)	5.008 (-2)	9.423 (-2)
	3.0	1.535 (-3)	4.551 (-3)	9.435 (-3)	1.670 (-2)	2.745 (-2)
	5.0	6.709 (-4)	1.851 (-3)	3.595 (-3)	5.860 (-3)	8.775 (-3)
96	0.5	4.890 (-2)	1.129 (-1)	2.713 (-1)	6.142 (-1)	1.341 (0)
	1.0	1.206 (-2)	3.706 (-2)	9.438 (-2)	2.156 (-1)	4.684 (-1)
	1.8	4.222 (-3)	1.383 (-2)	3.155 (-2)	6.216 (-2)	1.147 (-1)
	3.0	1.804 (-3)	5.799 (-3)	1.212 (-2)	2.130 (-2)	3.461 (-2)
	5.0	7.913 (-4)	2.370 (-3)	4.664 (-3)	7.589 (-3)	1.130 (-2)
Z	ω/m	$\xi_1^{(0)}$	$\xi_2^{(0)}$	$\xi_3^{(0)}$	$\xi_4^{(0)}$	$\xi_5^{(0)}$
10	0.3	4.465 (-4)	3.321 (-3)	2.461 (-2)	1.829 (-1)	1.264 (0)
	0.5	1.356 (-4)	6.769 (-4)	3.282 (-3)	1.608 (-2)	7.805 (-2)
	1.0	2.993 (-5)	9.132 (-5)	2.728 (-4)	8.105 (-4)	2.404 (-3)
	1.8	9.379 (-6)	2.051 (-5)	4.479 (-5)	9.239 (-5)	1.949 (-4)
	3.0	3.791 (-6)	6.585 (-6)	1.124 (-5)	1.895 (-5)	3.175 (-5)
20	0.3	4.313 (-3)	3.159 (-2)	2.265 (-1)	1.623 (0)	1.176 (1)
	0.5	1.245 (-3)	6.139 (-3)	2.940 (-2)	1.408 (-1)	6.737 (-1)
	1.0	2.584 (-4)	8.928 (-4)	2.385 (-3)	7.006 (-3)	2.952 (-2)
	1.8	7.685 (-5)	1.743 (-4)	3.755 (-4)	7.333 (-4)	1.662 (-3)
	3.0	2.965 (-5)	5.409 (-5)	9.415 (-5)	1.602 (-4)	2.912 (-4)
30	0.3	1.857 (-2)	1.311 (-1)	9.187 (-1)	6.292 (0)	4.322 (1)
	0.5	5.101 (-3)	2.503 (-2)	1.165 (-1)	5.400 (-1)	2.508 (0)
	1.0	9.937 (-4)	3.132 (-3)	9.204 (-3)	2.661 (-2)	7.663 (-2)
	1.8	2.739 (-4)	6.163 (-4)	1.420 (-3)	2.986 (-3)	6.206 (-3)
	3.0	1.024 (-4)	1.904 (-4)	3.477 (-4)	5.949 (-4)	9.599 (-4)
40	0.3	5.865 (-2)	4.141 (-1)	2.645 (0)	1.681 (1)	1.070 (2)
	0.5	1.544 (-2)	7.465 (-2)	3.329 (-1)	1.474 (0)	6.536 (0)
	1.0	2.834 (-3)	8.975 (-3)	2.585 (-2)	7.302 (-2)	2.032 (-1)
	1.8	7.510 (-4)	1.818 (-3)	3.929 (-3)	9.161 (-3)	1.675 (-2)
	3.0	2.614 (-4)	5.243 (-4)	9.405 (-4)	1.610 (-3)	2.606 (-3)

(Continued)

Z	ω/m	$\xi_1^{(0)}$	$\xi_2^{(0)}$	$\xi_3^{(0)}$	$\xi_4^{(0)}$	$\xi_5^{(0)}$
54	0.3	2.317 (-1)	1.520 (0)	8.376 (0)	4.518 (1)	2.429 (2)
	0.5	5.831 (-2)	2.672 (-1)	1.086 (0)	4.360 (0)	1.751 (1)
	1.0	9.973 (-3)	3.085 (-2)	8.461 (-2)	2.270 (-1)	6.059 (-1)
	1.8	2.467 (-3)	6.013 (-3)	1.709 (-2)	2.570 (-2)	5.123 (-2)
	3.0	7.506 (-4)	1.646 (-3)	2.984 (-3)	5.057 (-3)	8.345 (-3)
64	0.3	2.732 (-4)	5.100 (-4)	8.245 (-4)	1.237 (-3)	1.793 (-3)
	0.5	5.753 (-4)	3.441 (0)	1.613 (1)	7.280 (1)	3.257 (2)
	1.0	1.417 (-4)	6.019 (-1)	2.215 (0)	3.036 (0)	2.914 (1)
	1.8	2.332 (-2)	6.848 (-2)	1.777 (-1)	4.519 (-1)	1.145 (0)
	3.0	5.520 (-3)	1.311 (-2)	2.680 (-2)	5.259 (-2)	1.017 (-1)
72	0.3	1.079 (-3)	3.553 (-3)	6.271 (-3)	1.044 (-2)	1.093 (-2)
	0.5	5.401 (-4)	1.064 (-3)	1.706 (-3)	2.547 (-3)	3.664 (-3)
	1.0	1.186 (0)	6.344 (0)	2.488 (1)	9.504 (1)	3.349 (2)
	1.8	2.892 (-1)	1.116 (0)	3.697 (0)	1.206 (1)	3.930 (1)
	3.0	4.650 (-2)	1.268 (-1)	3.098 (-1)	7.451 (-1)	1.789 (0)
78	0.3	1.070 (-2)	2.411 (-2)	4.739 (-2)	8.991 (-2)	1.683 (-1)
	0.5	3.138 (-3)	6.456 (-3)	1.113 (-2)	1.814 (-2)	2.880 (-2)
	1.0	9.565 (-4)	1.895 (-3)	3.010 (-3)	4.444 (-3)	6.320 (-3)
	1.8	2.075 (0)	9.877 (0)	3.254 (1)	9.856 (1)	2.906 (2)
	3.0	5.030 (-1)	1.756 (0)	5.272 (0)	1.557 (1)	4.596 (1)
83	0.3	7.989 (-2)	2.005 (-1)	4.630 (-1)	1.058 (0)	2.421 (0)
	0.5	1.808 (-2)	3.812 (-2)	7.212 (-2)	1.325 (-1)	2.409 (-1)
	1.0	5.190 (-3)	1.016 (-2)	1.706 (-2)	2.721 (-2)	4.238 (-2)
	1.8	1.528 (-3)	2.952 (-3)	4.616 (-3)	6.724 (-3)	9.447 (-3)
	3.0	8.171 (-4)	2.556 (0)	6.957 (0)	1.862 (1)	4.979 (1)
88	0.3	1.289 (-1)	2.946 (-1)	6.430 (-1)	1.399 (0)	3.054 (0)
	0.5	2.888 (-2)	5.617 (-2)	1.022 (-1)	1.820 (-1)	3.217 (-1)
	1.0	8.179 (-3)	1.497 (-2)	2.441 (-2)	3.809 (-2)	5.821 (-2)
	1.8	2.354 (-3)	4.328 (-3)	6.630 (-3)	9.514 (-3)	1.320 (-2)
	3.0	1.373 (0)	3.726 (0)	9.029 (0)	2.149 (1)	5.111 (1)
92	0.3	2.156 (-1)	4.356 (-1)	8.500 (-1)	1.829 (0)	3.784 (0)
	0.5	4.801 (-2)	8.302 (-2)	1.452 (-1)	2.403 (-1)	4.265 (-1)
	1.0	1.349 (-2)	2.234 (-2)	3.512 (-2)	5.340 (-2)	7.982 (-2)
	1.8	3.824 (-3)	6.454 (-3)	9.613 (-3)	1.353 (-2)	1.848 (-2)
	3.0	2.152 (0)	5.037 (0)	1.089 (1)	2.339 (1)	4.976 (1)
96	0.3	3.379 (-1)	6.003 (-1)	1.153 (0)	2.250 (0)	4.435 (0)
	0.5	7.514 (-2)	1.162 (-1)	1.929 (-1)	3.201 (-1)	5.323 (-1)
	1.0	2.104 (-2)	3.119 (-2)	4.731 (-2)	7.020 (-2)	1.028 (-1)
	1.8	5.932 (-3)	9.028 (-3)	1.306 (-2)	1.805 (-2)	2.429 (-2)
	3.0	3.529 (0)	6.901 (0)	1.319 (1)	2.466 (1)	4.596 (1)

$$U_{1 \rightarrow f} = -\frac{e}{4\pi} (\alpha \alpha_{LM}^{(0)})_{21} \int_{r_1 > R} \psi_2^* B_{LM}^{(0)} \psi_1 dr_2, \quad (39.33)$$

and the Coulomb wave functions may be used in (39.33). Thus, the integrals for the R_1 should have a lower limit $r_2 = R$. The finite dimensions of the nucleus lead to corrections which are insignificant for small values of Z , but can be as great as 30-40 per cent when $Z \sim 80-90$.

§ 40. Conversion With Pair Creation. Nuclear Excitation by Electrons.

1. Conversion of Magnetic Multipole Radiation.

If the excitation energy ω of the nucleus is greater than $2m$, then in addition to conversion on an atomic shell, electron-positron pair creation may also take place. The conversion coefficient with pair creation can be found from the general formula (39.12), with ψ_2 an electron wave function, and ψ_1 a negative-frequency wave function corresponding to a positron. The matrix element entering (39.12) can be reduced to integrals of the radial wave functions, as was done in the previous paragraph for K -shell conversion.

We shall restrict our considerations to cases in which the electron and positron may be considered free particles¹⁾ (applicable to low- Z nuclei).

We shall choose the wave functions in the form of plane waves

$$\begin{aligned} \psi_2 &= u e^{-i\mathbf{p}_- \cdot \mathbf{r} - i\epsilon_- t}, \\ \psi_1 &= v e^{i\mathbf{p}_+ \cdot \mathbf{r} - i\epsilon_+ t}. \end{aligned} \quad (40.1)$$

Here $\mathbf{p}_-$ is the electron momentum, $\mathbf{p}_+$ is the positron momentum, and u and v are unit bispinor amplitudes.

When we insert (40.1) into (39.12) we obtain the following expression for the differential conversion coefficient of an electric or magnetic multipole with creation of an electron in the momentum interval $d\mathbf{p}_-$ and a positron in the momentum interval $d\mathbf{p}_+$:

$$d\beta_{LM}^{(0)} = \frac{e\omega}{4} \frac{d\mathbf{p}_- d\mathbf{p}_+}{(2\pi)^6} \delta(\omega - \epsilon_- - \epsilon_+) \Sigma \left| \int e^{i\mathbf{r} \cdot \mathbf{u}} B_{LM}^{(0)} \psi_1 dr \right|^2, \quad (40.2)$$

where

$$q = \mathbf{p}_- + \mathbf{p}_+$$

¹⁾ V. Berestetsky and I. Shmushkevich, J. Exptl.-Theoret. Phys. 12, 591 (1949); I. Shapiro, J. Exptl.-Theoret. Phys. 19, 597 (1949). M. Rev. Phys., Rev. 76, 678 (1949).

and Σ denotes summation over the electron and positron spin states.

Let us first perform the calculation for the case of a magnetic multipole. In this case, according to (39.10),

$$B_{LM}^{(0)} = -\alpha Y_{LM}^{(0)} G_L(\omega r).$$

Expanding $e^{i\mathbf{q} \cdot \mathbf{r}}$ in spherical functions, we arrive at

$$\int e^{i\mathbf{q} \cdot \mathbf{r}} Y_{LM}^{(0)} G_L(\omega r) dr = Y_{LM}^{(0)} \left(\frac{q}{\omega} \right) I_L,$$

where

$$I_L = \int G_L(\omega r) g_L(qr) r^2 dr = \frac{(4\pi)^{1/2}}{\omega(\omega^2 - q^2)^{1/2}} \left(\frac{q}{\omega} \right)^L.$$

Thus,

$$d\beta_{LM}^{(0)} = \frac{e}{\pi^2} \frac{1}{(\omega^2 - q^2)^{1/2}} \left(\frac{q}{\omega} \right)^{2L} d\mathbf{p}_+ d\mathbf{p}_- \Sigma \left| u^* \alpha Y_{LM}^{(0)} \left(\frac{q}{\omega} \right) v \right|^2 \delta(\omega - \epsilon_- - \epsilon_+). \quad (40.3)$$

The summation over spin states may be performed in the usual way. We have

$$\Sigma \left| u^* \alpha Y_{LM}^{(0)} v \right|^2 = \frac{1}{\epsilon_- \epsilon_+} \{ (\epsilon_- \epsilon_+ + m^2 - \mathbf{p}_- \cdot \mathbf{p}_+) | Y_{LM}^{(0)} |^2 + (\mathbf{p}_- \cdot \mathbf{Y}_{LM}^{(0)}) (\mathbf{p}_+ \cdot \mathbf{Y}_{LM}^{(0)}) + (\mathbf{p}_- \cdot \mathbf{Y}_{LM}^{(0)}) (\mathbf{p}_+ \cdot \mathbf{Y}_{LM}^{(0)}) \}.$$

This expression can be simplified. Since $\mathbf{p}_+ = \mathbf{q} - \mathbf{p}_-$ and $\mathbf{q} \cdot \mathbf{Y}_{LM}^{(0)} = 0$, we may write

$$\mathbf{p}_- \cdot \mathbf{Y}_{LM}^{(0)} = -\mathbf{p}_+ \cdot \mathbf{Y}_{LM}^{(0)}.$$

Further,

$$\mathbf{p}_- \cdot \mathbf{p}_+ = \mathbf{p}_+ \cdot \mathbf{q} - p_+^2.$$

On the other hand, conservation of energy

$$\omega = \sqrt{p_+^2 + m^2} + \sqrt{(q - p_-)^2 + m^2}$$

leads to

$$p_+ q = \omega \epsilon_+ - \frac{1}{2} (\omega^2 - q^2).$$

Thus,

$$\Sigma |u^* \alpha Y_{LM}^{(0)} v|^2 = \frac{1}{\epsilon_+ \epsilon_+} \left[\frac{\omega^2 - q^2}{2} |Y_{LM}^{(0)}|^2 - 2 |p_+ Y_{LM}^{(0)}|^2 \right].$$

Let us insert this expression into (40.3) and transform from the variables p_+ , p_- to the variables q , p_+ . We shall write the products of differentials in the form

$$dp_+ dp_- = dp_+ dq = p_+^2 dp_+ d\omega q^2 dq d\phi,$$

and the argument of the δ -function in (40.3) in the form

$$\omega - \epsilon_- - \epsilon_+ = \omega - \epsilon_+ - \sqrt{q^2 + p_-^2 + m^2} - 2q p_+ \cos \phi,$$

where ϕ is the angle between q and p_+ . If we choose the z axis along q , then

$$d\phi = d \cos \phi d\varphi,$$

and we can eliminate the δ -function according to

$$\delta(\omega - \epsilon_+ - \epsilon_-) d \cos \phi \rightarrow \frac{\epsilon_-}{p_+ q}.$$

The angle ϕ is now determined by the conservation of energy:

$$(\omega - \epsilon_+)^2 = q^2 + \epsilon_+^2 - 2q p_+ \cos \phi. \quad (40.4)$$

Equation (40.3) then becomes

$$d\beta_{LM}^{(0)} = \frac{a}{\pi^2 (\omega^2 - q^2)^2} \left(\frac{q}{\omega} \right)^{2L+1} \times \\ \times \left[\frac{\omega^2 - q^2}{2} |Y_{LM}^{(0)}|^2 - 2p_+^2 |Y_{LM}^{(0)}|^2 \sin^2 \phi \right] d\epsilon_+ dq d\varphi d\phi. \quad (40.5)$$

Let us integrate (40.5) over $d\varphi$ and $d\phi$. Taking the value of $\sin^2 \phi$ from Equation (40.4), we obtain the following expression for the conversion coefficient for pair creation with a positron energy ϵ_+ and a total momentum q :

$$d\beta_{LM}^{(0)}(\epsilon_+, q) = \frac{2a}{\pi} \frac{q^{2L-1}}{\omega^{2L+1}} \left\{ -\frac{1}{4} + \frac{p_+^2 + p_-^2}{2(\omega^2 - q^2)} + \frac{\omega^2}{(\omega^2 - q^2)^2} \right\} d\epsilon_+ dq. \quad (40.6)$$

Integrating this equation over q between the limits $q_{\min}$ and $q_{\max}$, where

$$q_{\min} = |p_+ - p_-|, \quad q_{\max} = |p_+ + p_-|,$$

we obtain the energy distribution of the conversion positrons:¹⁾

$$d\beta_{LM}^{(0)}(\epsilon_+) = \frac{a}{\pi \omega^{2L+1}} \left\{ \omega^{2(L-1)} [p_+^2 + p_-^2 - 2(L-1)m^2] \ln \frac{m^2 + p_+ p_- + \epsilon_+ \epsilon_-}{m\omega} - \right. \\ \left. - \frac{1}{4L} [(p_+ + p_-)^{2L} - (p_+ - p_-)^{2L}] + \frac{1}{2} [(\epsilon_+ \epsilon_- + p_+ p_- - m^2)(p_+ + p_-)^{2(L-1)} - \right. \\ \left. - (\epsilon_+ \epsilon_- - p_+ p_- - m^2)(p_+ - p_-)^{2(L-1)}] - \frac{1}{2} [p_+^2 + p_-^2 - 2(L-1)m^2] \times \right. \\ \left. \times \sum_{n=1}^{L-1} \frac{\omega^{2(L-n-1)}}{\pi} [(p_+ + p_-)^{2n} - (p_+ - p_-)^{2n}] \right\} d\epsilon_+. \quad (40.7)$$

¹⁾ When $L=0$ the last term in (40.7) which contains the sum over n from 1 to $L-1$ is set equal to zero.

We present explicit expressions for the cases of a magnetic dipole ($L = 1$), quadrupole ($L = 2$), and octupole ($L = 3$):

$$\left. \begin{aligned} d_{p_+}^{(1)}(e_+) &= \frac{a}{\pi\omega^3} (p_+^2 + p_-^2) \ln \frac{m^2 + p_+ p_- + \epsilon_+ \epsilon_-}{m\omega} d\epsilon_+; \\ d_{p_+}^{(2)}(e_+) &= \frac{a}{\pi\omega^5} \left\{ \omega^2 (p_+^2 + p_-^2 - 2m^2) \ln \frac{m^2 + p_+ p_- + \epsilon_+ \epsilon_-}{m\omega} + \right. \\ &\quad \left. + p_+ p_- (\omega^2 - 3p_+^2 - 3p_-^2) \right\} d\epsilon_+; \\ d_{p_+}^{(3)}(e_+) &= \frac{a}{\pi\omega^7} \left\{ \omega^4 (p_+^2 + p_-^2 - 4m^2) \ln \frac{m^2 + p_+ p_- + \epsilon_+ \epsilon_-}{m\omega} + \right. \\ &\quad \left. + \frac{2}{3} p_+ p_- [\omega^2 (8m^2 - 2\omega^2 (p_+^2 + p_-^2) - 5(p_+^2 + p_-^2)^2)] \right\} d\epsilon_+, \end{aligned} \right\} \quad (40.8)$$

The total conversion coefficient $\beta_L^{(s)}$ is obtained by integrating (40.7) from m to $\omega - m$. The results of this integration can not in general be expressed in terms of elementary functions. Later we shall present the results of a numerical integration. First let us investigate the ultrarelativistic case, $\omega \gg m$, when we may make the approximations $p_+ \approx \epsilon_+$, $p_- \approx \epsilon_-$. Assuming, in addition, that $\left(\frac{\omega}{m}\right)^2 \gg L$, we obtain

$$\beta_L^{(s)} = \frac{2\pi}{3\pi} \left[\ln \frac{2\omega}{m} - \frac{23}{12} - \frac{3}{4(2L+1)} - \sum_{n=1}^{L-1} \frac{4n+5}{2(2n+1)(2n+3)} \right]. \quad (40.9)$$

In the other limiting case $L \gg \left(\frac{\omega}{m}\right)^2$ (also with $\omega \gg m$), we find the following asymptotic expression for $\beta_L^{(s)}$:

$$\beta_L^{(s)} = \frac{aL}{\pi\omega^3} \left[\ln \frac{Lm^2}{\omega^2} + 2 + \gamma \right], \quad (40.10)$$

where γ is Euler's constant.

The distribution in the angle χ between the particles of the pair can be obtained by going from the variable q to χ in Equation (40.6), and integrating the resulting expression over ϵ_+ . The relation between χ and q is given by

$$q^2 = p_+^2 + p_-^2 - 2p_+ p_- \cos \chi.$$

As an example of an application of the formulas obtained for the conversion coefficient, let us determine the cross section for formation of a deuteron and creation of an electron-positron pair in slow neutron capture by a proton. Noting that in this process there is a transition from the 1_2 to the 3_2 state, we conclude that the transition is magnetic dipole. Therefore,

$$\sigma = \sigma_{ph} \beta_1^{(s)},$$

where σ is the cross section for the above process, and σ_{ph} is the photocapture cross section, which for slow neutrons is $\sigma_{ph} \approx 0.3 \cdot 10^{-24} \text{ cm}^2$. The deuteron binding energy is $| \epsilon | = 2.15 \text{ Mev}$, so that $\omega = 4.2 \text{ m}$. Numerical integration for this value of ω gives the value $3 \cdot 10^{-4}$ for $\beta_1^{(s)}$. Therefore,

$$\sigma = 0.9 \cdot 10^{-28} \text{ cm}^2.$$

2. Conversion of Electric-Multipole Radiation.

Let us now consider an electric multipole. Inserting the expression

$$-B_{LM}^{(1)} = \sqrt{\frac{L}{L+1}} G_L(\omega r) Y_{LM} + \frac{\sqrt{2L+1}}{\sqrt{L+1}} G_{L-1}(\omega r) Y_{L, L-1, M} \alpha$$

into (39.12), and again using the wave functions (40.1), we obtain

$$(B_{LM}^{(1)})^2 = \frac{1}{\omega(q^2 - \omega^2)} \left(\frac{q}{\omega} \right)^L \frac{1}{\sqrt{L+1}} \left[\sqrt{L} Y_{LM} u^* v - \sqrt{2L+1} \frac{\omega}{q} Y_{L, L-1, M} (u^* \alpha v) \right].$$

Using this expression in (39.12), and integrating over the angle between p_+ and q , we arrive at

$$d\beta_L^{(1)} = \frac{a^2 2^{L+1}}{\pi^2 \omega^{2L+1} (\omega^2 - q^2)^2 (L+1)} \sum \left| \sqrt{L} Y_{LM} u^* v - \sqrt{2L+1} \frac{\omega}{q} Y_{L, L-1, M} (u^* \alpha v) \right|^2 \times \\ \times \epsilon_+ \epsilon_- d\epsilon_+ d\epsilon_- d\varphi_+ d\varphi_-, \quad (40.11)$$

where Σ denotes summation over electron spins. The summation is

$$\sum_{|\dots|^2=L} |Y_{LM}|^2 (e_+ e_- + p_+ p_- - m^2) + (2L+1) \frac{\omega^2}{q^2} |Y_{L, L-1, M}|^2 \times \\ \times (e_+ e_- - p_+ p_- + m^2) - \sqrt{L(2L+1)} \frac{\omega}{q} (Y_{LM} Y_{L, L-1, M}^* + Y_{LM}^* Y_{L, L-1, M}) \times \\ \times (e_+ p_+ + e_- p_-) + (2L+1) \frac{\omega^2}{q^2} (Y_{L, L-1, M} p_-) (Y_{L, L-1, M}^* p_+) + \\ + (Y_{L, L-1, M}^* p_-) (Y_{L, L-1, M} p_+).$$

This scalar products ($\sum_{L, L-1, M} \mathbf{p}_+$) which enter into this expression are easily calculated by resolving into longitudinal and transverse spherical vectors. For the longitudinal vector we have

$$Y_{LM}^{(0)} \mathbf{p}_+ = p_+ Y_{LM} \cos \theta, \\ Y_{LM}^{(-1)} \mathbf{p}_- = (q - p_+ \cos \theta) Y_{LM},$$

and for the transverse one,

$$Y_{LM}^{(1)} \mathbf{p}_+ = -Y_{LM}^{(1)} \mathbf{p}_-, \\ |Y_{LM}^{(1)} \mathbf{p}_+|^2 = |Y_{LM}^{(1)}|^2 \sin^2 \theta.$$

Inserting these expressions into (40.11) and integrating over $d\varphi_+$ and $d\theta$, we obtain

$$d\beta_{\pm}^{(1)}(e_+, q) = \frac{2\pi}{\pi(L+1)} \left\{ \frac{q^{2L-1}}{\omega^{2L+1}} \left[\frac{L}{2} + \frac{(3L+1)\omega^2}{2} - 2Le_+ e_- \right] \frac{1}{\omega^2 - q^2} - \right. \\ \left. - \frac{2L\omega e_+ e_-}{(\omega^2 - q^2)^2} \right\} + \frac{q^{2L-3}}{\omega^{2L-1}} \left[\frac{3L+1}{3} - ((L+1)e_+ e_- + L\omega^2 + (L+1)m^2) \frac{1}{\omega^2 - q^2} + \right. \\ \left. + \omega^2 (2Le_+ e_- + (L+1)m^2) \frac{1}{(\omega^2 - q^2)^2} \right] d e_+ d\eta. \quad (40.12)$$

Integration over q leads to the energy distribution of the conversion positrons:

$$d\beta_{\pm}^{(1)}(e_+) = \frac{a}{\pi\omega^3} \left\{ \omega^2 (L-1) [e_+^2 + e_-^2 - 2(L-1)m^2] \ln \frac{m^2 + e_+ e_- + p_+ p_-}{m\omega} + \right. \\ \left. + \frac{1}{2(L-1)} [(p_+ + p_-)^{2L} - (p_+ - p_-)^{2L}] - \left[\frac{(e_+ - e_-)^2}{4(L+1)} + m^2 \right] [(p_+ + p_-)^{2(L-1)} - \right. \\ \left. - (p_+ - p_-)^{2(L-1)}] + \frac{\omega^2}{2} [(e_+ e_- + p_+ p_- - m^2)(p_+ + p_-)^{2(L-2)} - \right. \\ \left. - (e_+ e_- - p_+ p_- - m^2)(p_+ - p_-)^{2(L-2)}] - \frac{1}{2} [e_+^2 + e_-^2 - 2(L-1)m^2] \times \right. \\ \left. \times \sum_{n=1}^{L-1} \frac{\omega^{2(L-n-1)}}{n} [(p_+ + p_-)^{2n} - (p_+ - p_-)^{2n}] \right\} d e_+. \quad (40.13)$$

Equation (40.13) is valid for all L , except $L=1$. In the latter case (electric dipole), we have

$$d\beta_{\pm}^{(1)}(e_+) = \frac{a}{\pi\omega^3} \left\{ (e_+^2 + e_-^2) \ln \frac{m^2 + e_+ e_- + p_+ p_-}{m\omega} + 2p_+ p_- \right\} d e_+. \quad (40.14)$$

We present here the expressions for the electric quadrupole and octupole, as given by (40.13):

$$d\beta_{\pm}^{(2)}(e_+) = \frac{a}{\pi\omega^5} \left\{ \omega^2 (e_+^2 + e_-^2 - 2m^2) \ln \frac{m^2 + e_+ e_- + p_+ p_-}{m\omega} + \right. \\ \left. + \frac{8}{3} p_+ p_- (e_+ e_- - m^2) \right\} d e_+; \\ d\beta_{\pm}^{(3)}(e_+) = \frac{a}{\pi\omega^7} \left\{ \omega^4 (e_+^2 + e_-^2 - 4m^2) \ln \frac{m^2 + e_+ e_- + p_+ p_-}{m\omega} - \right. \\ \left. - p_+ p_- [2\omega^4 - \omega^2 m^2 - 8m^4 - (9\omega^2 + 4m^2) e_+ e_- + 4e_+^2 e_-^2] \right\} d e_+. \quad (40.15)$$

As for magnetic multipoles, we shall present the expression for the total internal conversion coefficient only in the limit of $\omega \gg m$. In this case, if $\left(\frac{\omega}{m}\right)^2 \gg L$, then

$$\beta_{\pm}^{(L)} = \frac{2\pi}{3\pi} \left\{ \ln \frac{2\omega}{m} - \frac{23}{12} + \frac{5L+1}{4(L+1)(2L+1)} - \sum_{n=1}^{L-1} \frac{4n+5}{2(2n+1)(2n+3)} \right\}. \quad (40.16)$$

This formula is valid for $L \geq 2$. When $L=1$,

$$\beta_{\pm}^{(1)} = \frac{2\pi}{3\pi} \left(\ln \frac{2\omega}{m} - \frac{5}{3} \right). \quad (40.17)$$

In the other limiting case when $\frac{L^2}{m} \gg \left(\frac{\omega}{m}\right)^2 \gg 1$, we obtain the same expression as (40.10) for the magnetic multipole.

We present below a table of values for the internal conversion coefficient with pair creation for electric and magnetic multipoles, as obtained by numerical integration of expressions (40.7) and (40.13).

3. Nuclear Excitation by Electrons.

The process which is the direct opposite of internal conversion of gamma rays is nuclear excitation by an electron. In this case the energy of the final electron state is lower than that of the initial one, $\epsilon_2 < \epsilon_1$, and the final energy of the nucleus is greater than the initial, $E_2 > E_1$. Using the results of Section 39 we may write down the expression for the probability of exciting a nucleus, changing its angular momentum by L , and its parity by $(-1)^{L+\lambda+1}$ together with the transition of an electron from an initial state (1) to a final state (2):

$$w_L^{(1)} = w_{L,rad}^{(1)} \beta_L^{(1)}, \quad (40.18)$$

where $\beta_L^{(1)}$ is the probability that the nucleus undergoes radiative transition from the excited (final) state to the initial one, and $\beta_L^{(1)}$ is the "conversion coefficient" given by

$$\beta_L^{(1)} = \frac{e^2}{4} \omega \left| \int \psi_2^* B_{L,M}^{(1)} \psi_1 d\mathbf{r} \right|^2. \quad (40.19)$$

As in treating conversion with pair creation, we shall restrict our considerations to the free-electron approximation.¹⁾ Let p_1 and p_2 be the electron momenta in the initial and final states, and let u_1 and u_2 be their corresponding plane-wave amplitudes. According to (40.18) and (40.19), the differential cross section for the process under consideration is

$$d\sigma_L^{(1)} = \frac{e^2}{4v_1} \omega_{L,rad}^{(1)} \delta(\epsilon_1 - \epsilon_2 - \omega) \frac{1}{2} \sum_{\lambda} \left| \int e^{iqr} u_2^* B_{L,M}^{(1)} u_1 d\mathbf{r} \right|^2, \quad (40.20)$$

where u_1 is the initial electron velocity,

$$q = p_0 - p_1$$

and Σ denotes summation over initial and final electron spins (the factor $\frac{1}{2}$ in front of the summation comes from averaging over initial spins). A comparison of (40.20) and (40.2) shows that the last factors in these equations, namely

¹⁾ G. Wick, *Ricerca sci. (Italy)* 11, 49 (1940); K. Ter-Martirosian, *J. Exptl.-Theoret. Phys.* 20, 925 (1950).

$$\Sigma \left| \int e^{i(p_1 - p_2)r} u_2^* B_{L,M}^{(1)} u_1 d\mathbf{r} \right|^2$$

in (40.20), and

$$\Sigma \left| \int e^{i(p_1 + p_2)r} u_2^* B_{L,M}^{(1)} u_1 d\mathbf{r} \right|^2$$

in (40.2), differ only by the substitution

$$\begin{aligned} p_- &\rightarrow p_1, \\ -p_+ &\rightarrow p_2, \\ -e_+ &\rightarrow e_2, \\ e_- &\rightarrow e_1. \end{aligned}$$

Therefore, (40.20) can be written

$$d\sigma_L^{(1)} = \frac{1}{2} \omega_{L,rad}^{(1)} (\epsilon_2 - \epsilon_1 - \omega) \frac{d^2 p_2}{v_1} \delta_L^{(1)}(p_1, -p_2), \quad (40.21)$$

where $\delta_L^{(1)}$ can be found from expressions (40.3), (40.11) for the differential conversion coefficients by writing them in the form

$$d\delta_L^{(1)} = \delta_L^{(1)}(p_-, p_+) \delta(\epsilon_+ + \epsilon_- - \omega) d\mathbf{p}_+ d\mathbf{p}_-.$$

Transforming from p_+ to q in (40.21), and integrating over angles, we obtain

$$d\sigma_L^{(1)} = \frac{1}{2} \omega_{L,rad}^{(1)} \delta_L^{(1)}(p_1, q) \frac{dq}{p_1^2}, \quad (40.22)$$

where

(Total Conversion Coefficients $\sigma_L^{(e)}$)

ω/m	3	4	5	6	7	10	15	20
1	$0.710 \cdot 10^{-4}$	$0.537 \cdot 10^{-3}$	$0.469 \cdot 10^{-3}$	$0.675 \cdot 10^{-3}$	$0.884 \cdot 10^{-3}$	$0.134 \cdot 10^{-2}$	$0.193 \cdot 10^{-2}$	$0.206 \cdot 10^{-2}$
2	$0.273 \cdot 10^{-4}$	$0.141 \cdot 10^{-3}$	$0.301 \cdot 10^{-3}$	$0.471 \cdot 10^{-3}$	$0.637 \cdot 10^{-3}$	$0.108 \cdot 10^{-2}$	$0.164 \cdot 10^{-2}$	$0.206 \cdot 10^{-2}$
3	$0.114 \cdot 10^{-4}$	$0.483 \cdot 10^{-4}$	$0.204 \cdot 10^{-3}$	$0.346 \cdot 10^{-3}$	$0.491 \cdot 10^{-3}$	$0.902 \cdot 10^{-3}$	$0.141 \cdot 10^{-2}$	$0.186 \cdot 10^{-2}$
4	$0.503 \cdot 10^{-5}$	$0.311 \cdot 10^{-4}$	$0.144 \cdot 10^{-3}$	$0.292 \cdot 10^{-3}$	$0.390 \cdot 10^{-3}$	$0.770 \cdot 10^{-3}$	$0.129 \cdot 10^{-2}$	$0.169 \cdot 10^{-2}$
5	$0.299 \cdot 10^{-5}$	$0.223 \cdot 10^{-4}$	$0.103 \cdot 10^{-3}$	$0.203 \cdot 10^{-3}$	$0.297 \cdot 10^{-3}$	$0.608 \cdot 10^{-3}$	$0.117 \cdot 10^{-2}$	$0.157 \cdot 10^{-2}$
10	$0.59 \cdot 10^{-7}$	$0.409 \cdot 10^{-5}$	$0.248 \cdot 10^{-4}$	$0.680 \cdot 10^{-4}$	$0.130 \cdot 10^{-3}$	$0.372 \cdot 10^{-3}$	$0.792 \cdot 10^{-3}$	$0.115 \cdot 10^{-2}$

Total Conversion Coefficients $\sigma_L^{(e)}$

ω/m	3	4	5	6	7	10	15	20
1	$0.291 \cdot 10^{-3}$	$0.671 \cdot 10^{-3}$	$0.998 \cdot 10^{-3}$	$0.127 \cdot 10^{-2}$	$0.170 \cdot 10^{-2}$	$0.205 \cdot 10^{-2}$	$0.267 \cdot 10^{-2}$	$0.312 \cdot 10^{-2}$
2	$0.100 \cdot 10^{-3}$	$0.333 \cdot 10^{-3}$	$0.581 \cdot 10^{-3}$	$0.811 \cdot 10^{-3}$	$0.101 \cdot 10^{-2}$	$0.152 \cdot 10^{-2}$	$0.212 \cdot 10^{-2}$	$0.256 \cdot 10^{-2}$
3	$0.393 \cdot 10^{-4}$	$0.152 \cdot 10^{-3}$	$0.373 \cdot 10^{-3}$	$0.565 \cdot 10^{-3}$	$0.746 \cdot 10^{-3}$	$0.121 \cdot 10^{-2}$	$0.179 \cdot 10^{-2}$	$0.222 \cdot 10^{-2}$
4	$0.165 \cdot 10^{-4}$	$0.109 \cdot 10^{-3}$	$0.253 \cdot 10^{-3}$	$0.413 \cdot 10^{-3}$	$0.573 \cdot 10^{-3}$	$0.100 \cdot 10^{-2}$	$0.156 \cdot 10^{-2}$	$0.198 \cdot 10^{-2}$
5	$0.122 \cdot 10^{-4}$	$0.088 \cdot 10^{-3}$	$0.177 \cdot 10^{-3}$	$0.312 \cdot 10^{-3}$	$0.485 \cdot 10^{-3}$	$0.856 \cdot 10^{-3}$	$0.129 \cdot 10^{-2}$	$0.180 \cdot 10^{-2}$
10	$0.17 \cdot 10^{-6}$	$0.784 \cdot 10^{-5}$	$0.393 \cdot 10^{-4}$	$0.978 \cdot 10^{-4}$	$0.174 \cdot 10^{-3}$	$0.451 \cdot 10^{-3}$	$0.900 \cdot 10^{-3}$	$0.127 \cdot 10^{-2}$

$$\sigma_L^{(e)}(p_+, q) = \frac{d\sigma_L^{(e)}(e_+, q)}{d\epsilon_+, dq}$$

and $\frac{d\sigma_L^{(e)}(\lambda)}{d\epsilon_+, dq}$ is given by (40.6) and (40.12).

In order to obtain the total cross section $\sigma_L^{(e)}$ for excitation of the nucleus, (40.22) must be integrated over q from $q_{\min}$ to $q_{\max}$ these limits being determined by the conservation laws. Since the total cross section cannot depend on the direction of the initial electron momentum, (40.22) can also be averaged over directions of the vector $\mathbf{p}_1$:

$$\sigma_L^{(e)} = \frac{1}{2} w_{\text{rad}}^{(e)} \frac{1}{p_1^2} \frac{1}{4\pi} \int \frac{d\Omega_1}{4\pi} \int_{q_{\min}}^{q_{\max}} \sigma_L^{(e)}(\lambda)(p_1, q) dq.$$

This expression can be written

$$\sigma_L^{(e)} = \frac{w_{\text{rad}}^{(e)}}{8\pi p_1^2} b_L^{(e)}(e_1), \quad (40.23)$$

where

$$b_L^{(e)}(e) = \frac{d\sigma_L^{(e)}(e)}{de}$$

and $\frac{d\sigma_L^{(e)}(\lambda)}{d\epsilon_+, dq}$ is given by (40.7) and (40.13) with p_+ and p_- replaced by p_1 and p_2 .

4. Monochromatic Positrons.

If an atomic shell has an unfilled state, then it is possible to produce a pair in which the electron will occupy this bound state, and the positron will have a well-defined and definite energy. This is the phenomenon of internal conversion with emission of monochromatic positrons. In calculating the internal conversion coefficient in this case, Equation (39.13) can be used directly with ψ_2 a wave function of the discrete electron spectrum, and ψ_1 a negative-frequency wave function of the continuous spectrum, normalized for a unit energy interval. Equations (39.25) and (39.28) are then valid for the K-shell, and the radial functions of the continuous spectrum g_{-k} and f_{-k} belong to the negative frequencies. In the approximation of low Z and energies which are high compared with the K-shell binding energy, we obtain

(Total Conversion Coefficients $\beta_L^{(a)}$)

L	w/π	3	4	5	6	7	10	15	20
1	$0.710 \cdot 10^{-4}$	$0.237 \cdot 10^{-3}$	$0.469 \cdot 10^{-3}$	$0.675 \cdot 10^{-3}$	$0.854 \cdot 10^{-3}$	$0.134 \cdot 10^{-2}$	$0.193 \cdot 10^{-2}$	$0.236 \cdot 10^{-2}$	$0.266 \cdot 10^{-2}$
2	$0.273 \cdot 10^{-4}$	$0.141 \cdot 10^{-3}$	$0.301 \cdot 10^{-3}$	$0.471 \cdot 10^{-3}$	$0.637 \cdot 10^{-3}$	$0.108 \cdot 10^{-2}$	$0.154 \cdot 10^{-2}$	$0.186 \cdot 10^{-2}$	$0.206 \cdot 10^{-2}$
3	$0.114 \cdot 10^{-4}$	$0.833 \cdot 10^{-4}$	$0.204 \cdot 10^{-3}$	$0.346 \cdot 10^{-3}$	$0.491 \cdot 10^{-3}$	$0.502 \cdot 10^{-3}$	$0.41 \cdot 10^{-3}$	$0.35 \cdot 10^{-3}$	$0.30 \cdot 10^{-3}$
4	$0.503 \cdot 10^{-5}$	$0.311 \cdot 10^{-4}$	$0.144 \cdot 10^{-3}$	$0.282 \cdot 10^{-3}$	$0.350 \cdot 10^{-3}$	$0.770 \cdot 10^{-3}$	$0.129 \cdot 10^{-2}$	$0.169 \cdot 10^{-2}$	$0.199 \cdot 10^{-2}$
5	$0.229 \cdot 10^{-5}$	$0.233 \cdot 10^{-4}$	$0.103 \cdot 10^{-3}$	$0.203 \cdot 20^{-3}$	$0.297 \cdot 10^{-3}$	$0.608 \cdot 10^{-3}$	$0.117 \cdot 10^{-2}$	$0.157 \cdot 10^{-2}$	$0.197 \cdot 10^{-2}$
10	$0.53 \cdot 10^{-7}$	$0.409 \cdot 10^{-5}$	$0.248 \cdot 10^{-4}$	$0.680 \cdot 10^{-4}$	$0.130 \cdot 10^{-3}$	$0.372 \cdot 10^{-3}$	$0.392 \cdot 10^{-3}$	$0.392 \cdot 10^{-3}$	$0.392 \cdot 10^{-3}$

(Total Conversion Coefficients $\beta_L^{(b)}$)

L	w/π	3	4	5	6	7	10	15	20
1	$0.291 \cdot 10^{-4}$	$0.071 \cdot 10^{-3}$	$0.095 \cdot 10^{-3}$	$0.127 \cdot 10^{-3}$	$0.170 \cdot 10^{-3}$	$0.170 \cdot 10^{-3}$	$0.205 \cdot 10^{-3}$	$0.257 \cdot 10^{-3}$	$0.312 \cdot 10^{-3}$
2	$0.100 \cdot 10^{-4}$	$0.333 \cdot 10^{-4}$	$0.381 \cdot 10^{-3}$	$0.811 \cdot 10^{-3}$	$0.101 \cdot 10^{-2}$	$0.192 \cdot 10^{-2}$	$0.312 \cdot 10^{-2}$	$0.396 \cdot 10^{-2}$	$0.466 \cdot 10^{-2}$
3	$0.393 \cdot 10^{-4}$	$0.185 \cdot 10^{-3}$	$0.373 \cdot 10^{-3}$	$0.565 \cdot 10^{-3}$	$0.746 \cdot 10^{-3}$	$0.121 \cdot 10^{-2}$	$0.179 \cdot 10^{-2}$	$0.222 \cdot 10^{-2}$	$0.272 \cdot 10^{-2}$
4	$0.165 \cdot 10^{-4}$	$0.109 \cdot 10^{-3}$	$0.232 \cdot 10^{-3}$	$0.413 \cdot 10^{-3}$	$0.573 \cdot 10^{-3}$	$0.100 \cdot 10^{-2}$	$0.156 \cdot 10^{-2}$	$0.198 \cdot 10^{-2}$	$0.248 \cdot 10^{-2}$
5	$0.732 \cdot 10^{-5}$	$0.068 \cdot 10^{-4}$	$0.177 \cdot 10^{-3}$	$0.312 \cdot 10^{-3}$	$0.485 \cdot 10^{-3}$	$0.856 \cdot 10^{-3}$	$0.159 \cdot 10^{-2}$	$0.300 \cdot 10^{-2}$	$0.466 \cdot 10^{-2}$
10	$0.17 \cdot 10^{-6}$	$0.784 \cdot 10^{-5}$	$0.393 \cdot 10^{-4}$	$0.978 \cdot 10^{-4}$	$0.174 \cdot 10^{-3}$	$0.451 \cdot 10^{-3}$	$0.900 \cdot 10^{-3}$	$0.900 \cdot 10^{-3}$	$0.900 \cdot 10^{-3}$

$$\beta_L^{(b)}(p_4, q) = \frac{d\beta_L^{(b)}(e_+, q)}{de_+, dq}$$

and $d\beta_L^{(b)}$ is given by (40.6) and (40.12).

In order to obtain the total cross section $\sigma_L^{(b)}$ for excitation of the nucleus, (40.22) must be integrated over q from $q_{\min}$ to $q_{\max}$ these limits being determined by the conservation laws. Since the total cross section cannot depend on the direction of the initial electron momentum, (40.22) can also be averaged over directions of the vector $\mathbf{p}_1$:

$$\sigma_L^{(b)} = \frac{1}{2} w_{\text{rad}}^{(b)} \frac{1}{p_1^2} \frac{1}{4\pi} \int \frac{d\Omega_1}{4\pi} \int_{q_{\min}}^{q_{\max}} \beta_L^{(b)}(p_1, q) dq.$$

This expression can be written

$$\sigma_L^{(b)} = \frac{w_{\text{rad}}^{(b)}}{8\pi p_1^2} b_L^{(b)}(e_1), \quad (40.23)$$

where

$$b_L^{(b)}(e) = \frac{d\beta_L^{(b)}(e)}{de}$$

and $d\beta_L^{(b)}$ is given by (40.7) and (40.13) with p_+ and p_- replaced by p_1 and p_2 .

4. Monochromatic Positrons.

If an atomic shell has an unfilled state, then it is possible to produce a pair in which the electron will occupy this bound state, and the positron will have a well-defined definite energy. This is the phenomenon of internal conversion with emission of monochromatic positrons. In calculating the internal conversion coefficient in this case, Equation (39.13) can be used directly with ψ_2 a wave function of the discrete electron spectrum, and ψ_1 a negative-frequency wave function of the continuous spectrum, normalized for a unit energy interval. Equations (39.26) and (39.28) are then valid for the K-shell, and the radial functions of the continuous spectrum p_K and f_{-K} belong to the negative frequencies. In the approximation of low Z and energies which are high compared with the K-shell binding energy, we obtain

$$\beta_P^{(1)} = 2\alpha(Z\alpha)^3 \frac{(L+1)\omega^2 + 4Lm^2}{(L+1)\omega^2} \left(1 - \frac{2m}{\omega}\right)^{L-1/2},$$

$$\beta_P^{(2)} = 2\alpha(Z\alpha)^3 \left(1 - \frac{2m}{\omega}\right)^{L+1/2}.$$

As a numerical example, we present the ratio between the probability for conversion with emission of monochromatic positrons (with creation of a K-electron) and the total probability for conversion with pair creation.¹⁾ According to calculations performed with the exact radial wave functions for $\omega = 1.4$ Mev, this ratio for an electric dipole is $1/2$. As ω decreases, this ratio increases.

The process which is the inverse of this one is the following. A positron collides with an electron in the K-shell of an atom, and this pair is absorbed by the nucleus, so that no photons are observed and the nucleus is excited. The probability for this process can be written

$$w = w_P \beta, \quad (40.24)$$

where β is the coefficient for internal conversion with creation of monochromatic positrons, and w_P is the probability for nuclear excitation by a photon of the appropriate energy.

Since the energy the nucleus acquires in this process is quite large, fission may result. For instance, the cross section for uranium fission by this pair-absorption process is of the order of 10^{-31} cm^2 .²⁾

5. Pair Production in Particle Collisions.

A process similar to internal conversion with pair creation is pair creation on collision of two charged particles. This process can be represented by the diagrams shown in Fig. 50.

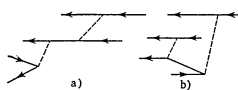


Fig. 50

If one of the particles can be replaced by an equivalent external field, then the upper lines of Fig. 50 may be removed. The diagrams so obtained are shown in Fig. 51, where the external dotted line represents the external field. We note that if the external field is accounted for in the electron wave functions, the diagrams of Fig. 51 are transformed to the internal conversion diagrams of Fig. 47.

¹⁾ L. Sliv, Proc. Acad. Sci. USSR 64, 321 (1949).

²⁾ R. Present and S. Chow, Phys. Rev. 85, 447 (1952).

Using Fig. 51 it is easy to construct the matrix element for pair creation by an electron in the field of the nucleus. We present here only an approximate expression for the cross section of this process in the ultra-relativistic case ($\epsilon \gg m$), where ϵ is the initial electron energy.³⁾

$$\sigma \cong \frac{r_0^2 Z^2 \alpha^3}{\pi} \left(\ln \frac{\epsilon}{m} \right)^3, \quad (40.25)$$

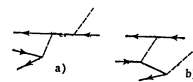


Fig. 51

We note that when $Z = 1$, pair production in collision of two electrons will be the same order of magnitude in this limiting case. The cross section of (40.25) has a coefficient which is $\frac{\alpha}{4\pi}$ times less than that of the bremsstrahlung cross section (31.20); the power of the logarithm, however, is in this case higher [Equation (40.25) does not take account of screening].

Figure 51 can also be used when considering the process of pair creation in collision of two heavy particles moving at a nonrelativistic velocity, since a system of two nonrelativistic particles is equivalent to a single particle in an external field. This process is equivalent to conversion with pair creation for an electric dipole (in this case the particle states belong to the continuous spectrum). An approximate expression for this cross section is⁴⁾

$$\sigma = r_0^2 \alpha^3 \frac{m^2 Z_1^2 Z_2^2}{M_2 T_2} \left(\frac{Z_1 M_2 - Z_2 M_1}{M_1} \right)^2 \quad (40.26)$$

(M_1 and M_2 are the masses of the particles; the particle with mass M_1 is at rest, and that with mass M_2 has a kinetic energy T_2 , which is assumed much greater than $2m$).

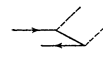


Fig. 52

It is also possible that both colliding particles can be replaced by external fields, so that the diagram illustrating the process degenerates to that shown in Fig. 52. In this diagram both dotted lines correspond to an external field. Its structure is the same as that for pair creation by two photons. This case occurs in collision of very high-energy particles, so that after pair creation their states of motion may be considered unchanged⁵⁾ (i.e., with no change of direction or velocity). The matrix element for this process is

$$\langle S | \hat{r} | r \rangle = e^2 \hat{u} \hat{a}_\mu \frac{\hat{f}^2}{f^2 + m^2} \hat{a}_\mu \hat{u},$$

³⁾ See W. Heitler, The Quantum Theory of Radiation (State Tech. Press, 1940) p. 220.

⁴⁾ L. Landau and E. Lifshits, Physik. Z. Sowjetunion 6, 244 (1934).

where $a_{\pm 1}$ and $a_{\pm 2}$ are the Fourier components of the external field which is a superposition of the fields of two charges moving at equal velocities (we note that it is just the superposition of the field that leads to pair production, a single free particle, naturally, cannot create pairs). The expression for the cross section for the case in which one of the particles is at rest is the same as (40.25), except that the nuclear charge should be replaced by the product of the charges $Z = Z_1 Z_2$.

Any radioactive decay with sufficient energy can be accompanied by pair production. This process can also be treated as pair production by an external field having frequencies greater than $2m$. Thus, when a nucleus disintegrates into two parts, the probability is given by¹⁾

$$w = \frac{2Z'^2 \alpha}{9\pi^2} \sigma^2 \ln \frac{2E}{M} \left(\ln \frac{2E}{m} - \frac{10}{3} \right), \quad (40.27)$$

where $Z' = Z_1 (1 - \frac{Z_2 A_2}{Z_1 A_1})$, A_1 and A_2 are the atomic masses of the fragments, E and v are the energy and velocity of the smaller fragment, and M is its mass. The probability for pair creation in β -decay is about²⁾ $10^{-8} - 10^{-7}$.

§ 41. 0 - 0 Transitions.

1. Reduction to the Static Interaction.

The general expressions derived in Section 39 for the probabilities of conversion or nuclear excitation become invalid if the angular momenta of the initial and final nuclear states are zero. When $L = 0$ Equation (39.11) vanishes, in agreement with the fact that there exists no photon state with angular momentum zero.

Nevertheless the element of $\delta^{(2)}$ for transitions between two such states is nonzero. Indeed, as has already been shown in Section 39, the integrals in (39.11) are taken over the region outside the nucleus, whereas the general expression for the matrix element, according to (39.32), contains an integral also over the region occupied by the nucleus, and this integral is in general different from zero when $L = 0$. Thus, the matrix element for the effective interaction energy is of the form

$$U_{i \rightarrow f} = \int_{\Omega} j_{21}(r_1) j_{21}(r_2) \frac{e^{i\omega |r_1 - r_2|}}{4\pi |r_1 - r_2|} dr_1 dr_2,$$

where j_{21} is the four-dimensional electron transition current, and J_{21} is the four-dimensional nuclear transition current. The integral is taken over the region Ω occupied by the nucleus. This matrix element can also be written in the form of (35.16), namely,

$$U_{i \rightarrow f} = \int_{\Omega} j_{21}(r_1) A_{21}(r_1) dr_1, \quad (41.1)$$

¹⁾ A. Migdal, Bull. Acad. Sci. USSR 4, 2, 287 (1940).

²⁾ L. Tissa, J. Exptl.-Theoret. Phys. 7, 690 (1937); E. Feinberg, Proc. Acad. Sci. USSR 23, 778 (1939).

where

$$A_{21}(r_1) = \int j_{21}(r_2) \frac{e^{i\omega |r_1 - r_2|}}{4\pi |r_1 - r_2|} dr_2. \quad (41.2)$$

The case $L = 0$ corresponds to spherical symmetry of a nuclear transition current distribution. Therefore, its space (J_{21}) and time (j_{21}^0) components should be of the form

$$J_{21}(r_2) = f(r_2) r_{2\mu}, \\ j_{21}^0(r_2) = g(r_2).$$

The potentials related to these currents have a similar structure:

$$A_{21}(r_2) = a_0(r_2) r_2 = -\nabla \lambda(r_2), \\ A_{21}^0(r_2) = \varphi_0(r_2). \quad (41.3)$$

[This is easily seen by using (39.2) in (41.2).]

Inserting (41.3) into (41.1) we obtain

$$U_{i \rightarrow f} = \int (j_{21}^0 \varphi_0 + j_{21} \nabla \lambda) dr. \quad (41.4)$$

Introducing the new function φ by the relation

$$\varphi_0 = \varphi - i\omega \lambda,$$

integrating the term containing $\nabla \lambda$ by parts, and using the continuity equation

$$\text{div } j_{21} = -i\omega j_{21}^0,$$

we obtain

$$U_{i \rightarrow f} = \int j_{21}^0 \varphi dr = e \int \psi_2^* \varphi \psi_1 dr. \quad (41.5)$$

We note that transformation of (41.4) to (41.5) is possible only because the general expression (41.1) for the matrix element is gauge invariant.

Let us determine the potential φ by making use of the fact that φ_0 satisfies the equation

$$\Delta \varphi_0 + \omega^2 \varphi_0 = -g(r),$$

and, therefore,

$$\Delta \varphi = \Delta (\varphi_0 + i\omega \lambda) = -g - \omega^2 \varphi_0 + i\omega \Delta \lambda.$$

Since $\underline{A} = -\nabla \lambda$, it follows from the Lorentz condition for the potentials that

$$\Delta \lambda = -i\omega \varphi_0$$

Therefore,

$$\Delta \varphi = -g = -e \Psi_2^* \Psi_1. \quad (41.5)$$

Equations (41.5) and (41.6) can also be written in the form

$$U_{i \rightarrow f} = e \int \Psi_2^*(r_1) \Psi_2(r_2) \Psi_1(r_1) \Psi_1(r_2) \frac{dr_1 dr_2}{|r_1 - r_2|}. \quad (41.7)$$

We note that when $\underline{L} = 0$ the transition is caused by the electrostatic (Coulomb) interaction of the charges. Equation (41.5) is conveniently written in the form

$$U_{i \rightarrow f} = \int J_{21}^0 \Phi dr, \quad (41.8)$$

where Φ is the electrostatic potential of the electron distribution, and satisfies the equation

$$\Delta \Phi = -e \Psi_2^* \Psi_1. \quad (41.9)$$

In this equation the right side may not have spherical symmetry. In view, however, of the spherical symmetry of the charge density in (41.8), a nonzero contribution is obtained only from the symmetric part of the potential. Therefore, the right side of (41.9) may be averaged over angles.

2. Conversion and Nuclear Excitation in 0-0 Transitions.

Let us find the probability for conversion and for nuclear excitation in 0-0 transitions.¹⁾ The electron wave function hardly changes in the region occupied by the nucleus. Therefore, the right side of (41.9) may be considered constant, and we may use the equation

$$\frac{1}{r} \frac{d^2(r\Phi)}{dr^2} = -e \rho_0,$$

where

$$\rho_0 = \Psi_2^*(0) \Psi_1(0).$$

The general solution of this equation is

$$\Phi = c_1 + \frac{c_2}{r} - e \rho_0 \frac{r^2}{6}.$$

The constant c_1 is of no significance, since (41.8) vanishes for constant Φ in view of the orthogonality of the nuclear wave functions. The constant c_2 is zero, which follows from the fact that Φ is finite when $r = 0$. Thus,

$$U_{i \rightarrow f} = -\frac{2\pi}{3} \rho_0 Q_0, \quad (41.10)$$

where

$$Q_0 = \frac{e^2}{4\pi} \int \Psi_2^* \Psi_1 r^2 dr. \quad (41.11)$$

The quantity Q_0 is the "zero-pole moment" of the nucleus, in analogy with the quadrupole moment. Since when $\underline{L} = 0$ no photon is radiated, we cannot introduce a conversion coefficient. To order of magnitude,

$$Q_0 \sim R^2 e^2,$$

¹⁾ Yukawa and Sakata, Proc. Phys.-Math. Soc. Japan 17, 10 (1935); K. Ter-Martirosian, J. Exptl.-Theoret. Phys. 20, 925 (1950).

where R is the radius of the nucleus.

The probability for conversion on the K-shell is thus given in the following way:

$$w = 2 \cdot 2\pi \left(\frac{2\pi}{3}\right) |Q_0|^2 |\psi_K(0)|^2 |\psi_\epsilon(0)|^2, \quad (41.12)$$

where $\psi_K(0)$ is the value of the K-electron wave function in the nucleus, and $\psi_\epsilon(0)$ is the value, also in the nucleus, of the continuous-spectrum electron wave function normalized for a unit energy interval, and corresponding to the energy $\epsilon = \omega + |\epsilon_K|$ and the angular momentum $\underline{l} = \frac{1}{2}$.

The probability for conversion with pair production is

$$dw = (2\pi) \frac{2\pi}{3} |Q_0|^2 \Sigma |\rho_0|^2 \frac{d\rho_+ d\rho_-}{(2\pi)^2} \delta(\omega - \epsilon_- - \epsilon_+),$$

In the free-electron approximation,

$$\rho_0 = \bar{u} v,$$

where $\bar{u}$ and v are the positive- and negative-frequency plane-wave amplitudes, and Σ denotes summation over electron and positron spins. We may perform the summation, obtaining

$$\Sigma |\rho_0|^2 = \frac{1}{4\epsilon_+ \epsilon_-} \text{Sp} \hat{\beta} (\hat{p}_- - m) \hat{\beta} (-\hat{p}_+ - m) = \frac{-p_- p_+ + \epsilon_- \epsilon_+ + m^2}{\epsilon_+ \epsilon_-},$$

so that

$$dw = \frac{1}{9} |Q_0|^2 |p_-| |p_+| d\Omega_+ d\Omega_- (\epsilon_- \epsilon_+ + m^2 - p_- p_+), \quad (41.13)$$

Similarly, the differential cross section for nuclear excitation is given by

$$d\sigma = \frac{1}{18} |Q_0|^2 \frac{|p_2|}{|p_1|} d\Omega_2 (\epsilon_1 \epsilon_2 + m^2 + p_1 p_2), \quad (41.14)$$

The total cross section for nuclear excitation is

$$\sigma = \frac{2\pi}{9} |Q_0|^2 \frac{|p_2|}{|p_1|} (\epsilon_1 \epsilon_2 + m^2). \quad (41.15)$$

CHAPTER VIII

RADIATIVE CORRECTIONS. VACUUM POLARIZATION.

§ 42. Third-Order S Matrix

1. Third-Order Matrix Elements

Let us now go on to the consideration of concrete physical effects due to the interaction of an electron with the zero-point oscillations of the electromagnetic field, as well as those due to the interaction of an external electromagnetic field with the zero-point oscillations of the electron-positron field.

We shall start by considering radiative transitions to electron scattering in an external field. Since the scattering is a first-order process, the effects in which we are interested arise in third-order perturbation theory; we shall restrict our considerations to the third order. Diagrams illustrating such third order effects are shown in Fig. 53. $A_\mu^{(0)}$ denotes the external electromagnetic field.

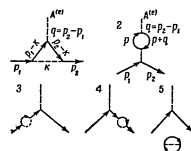


Fig. 53

Our previous results can be used immediately to exclude diagrams 3, 4, and 5, since 3 and 4 contain electron self-energy parts along the free-electron path, and 5 represents two independent processes one of which (corresponding to the closed electron loop) is a correction to the vacuum-vacuum transition probability which must be considered zero for physical reasons.

Thus the third-order processes which determine the interaction of the electron with the zero-point oscillations of the electromagnetic field as well as the interaction of the external electromagnetic field with the zero-point oscillations of the electron-positron field, correspond only to diagrams 1 and 2. The first of these will be called a

vertex diagram (radiative correction proper), and the second, a vacuum polarization diagram. According to (24.7), the matrix elements corresponding to these diagrams are

$$M_1^{(3)} = (-i)^3 e^3 u_2 \left(\int \gamma_\mu \frac{i(\hat{p}_2 - \hat{k}) - m}{(p_2 - k)^2 + m^2} \frac{\hat{a}(q)}{(2\pi)^4} \frac{i(\hat{p}_1 - \hat{k}) - m}{(p_1 - k')^2 + m^2} \gamma_\mu \frac{d^4 k}{k^2} \right) u_1, \quad (42.1)$$

$$M_2^{(3)} = -(-i)^3 e^3 u_2 \left(\int \gamma_\mu \frac{1}{(p + q)^2 + m^2} \text{Sp} \left\{ \frac{i(\hat{p} + \hat{x}) - m}{(p + q)^2 + m^2} \frac{\hat{a}(q)}{(2\pi)^4} \frac{i\hat{p} - m}{p^2 + m^2} \gamma_\mu \right\} d^4 p \right) u_1, \quad (42.2)$$

where $\hat{a}_\mu(q)$ is the Fourier transform of the external potential $A_\mu^{(0)}(x)$:

$$a_\mu(q) = \int A_\mu^{(0)}(x) e^{-iqx} d^4 x,$$

and u_1 and u_2 are the spinor amplitudes for the electron states with momenta p_1 and p_2 ; in addition, $q = p_2 - p_1$. The extra factor of -1 in (42.2), as compared with (42.1), is due to the fact that diagram 2 contains a closed electron loop. This is also the reason that (42.2) involves the trace Sp of the matrices belonging to this closed loop (see Section 24).

2. Calculation of the Matrix Element for Radiative Corrections to Electron Scattering¹⁾

Let us now calculate $M_1^{(3)}$. We shall write

$$M_1^{(3)} = \frac{ie^3}{(2\pi)^4} u_2 \mathfrak{U} u_1, \quad (42.3)$$

where

$$\begin{aligned} \mathfrak{U} &= \int \gamma_\mu \frac{i(\hat{p}_2 - \hat{k}) - m}{(p_2 - k)^2 + m^2} \hat{a}(q) \frac{i(\hat{p}_1 - \hat{k}) - m}{(p_1 - k')^2 + m^2} \gamma_\mu \frac{d^4 k}{k^2} = \\ &= \gamma_\mu (\hat{p}_2 - m) \hat{a}(q) (i\hat{p}_1 - m) \gamma_\mu J - i \gamma_\mu (\hat{p}_2 - m) \hat{a}(q) \gamma_\mu \gamma_\nu + \\ &\quad + \gamma_\mu \gamma_\nu \hat{a}(q) (i\hat{p}_1 - m) \gamma_\nu J - \gamma_\mu \gamma_\nu \hat{a}(q) \gamma_\nu \gamma_\mu J, \end{aligned} \quad (42.4)$$

and

$$\begin{aligned} J &= \int \frac{d^4 k}{(k^2 - 2p_1 k)(k^2 - 2p_2 k)k^2}, \\ J_\nu &= \int \frac{k_\nu d^4 k}{(k^2 - 2p_1 k)(k^2 - 2p_2 k)k^2}, \\ J_{\nu\sigma} &= \int \frac{k_\nu k_\sigma d^4 k}{(k^2 - 2p_1 k)(k^2 - 2p_2 k)k^2}. \end{aligned} \quad (42.5)$$

Only the third of these integrals is divergent for large $|k|$; according to the classification of divergences in Section 25, this integral is logarithmically divergent (diagram 1 is an irreducible vertex part). Therefore in the future we shall have to renormalize $J_{\sigma\tau}$, by subtracting its value for $q = 0$, writing

$$J_{\sigma\tau} = J_{\sigma\tau} - J_{\sigma\tau|q=0} \quad (42.6)$$

[see (26.11); in this case p_1 and p_2 are free-electron momenta].

We note that J diverges for small $|k|$. This represents the infrared catastrophe due to the inapplicability of perturbation theory in the low $|k|$ region (see Section 32).

Let us now use Equations (28.12) to simplify the operators multiplying the integrals J , J_ν , $J_{\sigma\tau}$ in (42.4).

The operator multiplying J can be written

$$\begin{aligned} \gamma_\mu (\hat{p}_2 - m) \hat{a}(q) (i\hat{p}_1 - m) \gamma_\mu &= \\ &= 2\hat{p}_1 \hat{a}(q) \hat{p}_2 - 2im [\hat{a}(q) \hat{p}_2 + \hat{p}_2 \hat{a}(q)] - 2im [\hat{a}(q) \hat{p}_1 + \hat{p}_1 \hat{a}(q)] - 2m^2 \hat{a}(q). \end{aligned} \quad (42.7)$$

¹⁾ See R. Feynman, Phys. Rev. 76, 769 (1949).

Since this expression is multiplied on the left by $\hat{u}_2$ and on the right by $\hat{u}_1$ in the matrix element, it is clear that the operator $\hat{p}_1$ on the right can be replaced by $-\hat{m}$. In the same way, $-\hat{m}$ can be substituted for $\hat{p}_2$ on the left. (If $\hat{p}_1$, rather than $\hat{p}_2$ is on the left, or if $\hat{p}_2$, rather than $\hat{p}_1$, is on the right, then before performing this substitution we must write $\hat{p}_1 = \hat{p}_2 - \hat{q}$, $\hat{p}_2 = \hat{p}_1 + \hat{q}$). After performing these substitutions we may rewrite (42.7) in the form

$$\gamma_\mu (\hat{p}_2 - m) \hat{a}(q) (\hat{p}_1 - m) \gamma_\mu = 4m^2 \hat{a}(q) - 2\hat{q} \hat{a}(q) \hat{q}. \quad (42.8)$$

Let us now make use of the relation¹⁾

$$\hat{q} \hat{a}(q) \hat{q} = -q^2 \hat{a}(q). \quad (42.8)$$

Then the final expression for the operator multiplying $\underline{J}$ is

$$\gamma_\mu (\hat{p}_2 - m) \hat{a}(q) (\hat{p}_1 - m) \gamma_\mu = 4m^2 \hat{a}(q) + 2q^2 \hat{a}(q). \quad (42.9)$$

We may similarly transform the operators multiplying $\underline{J}_0$ and $\underline{J}_{0\tau}$ in (42.4), obtaining

$$\gamma_\mu (\hat{p}_2 - m) \hat{a}(q) \gamma_\mu \gamma_\nu + \gamma_\mu \gamma_\nu \hat{a}(q) (\hat{p}_1 - m) \gamma_\mu = -4ma_\nu(q) + 2i[\hat{q} \hat{a}(q) \gamma_\nu - \gamma_\nu \hat{a}(q) \hat{q}], \quad (42.9')$$

$$\gamma_\mu \gamma_\nu \hat{a}(q) \gamma_\mu \gamma_\nu = -2\gamma_\nu \hat{a}(q) \gamma_\nu. \quad (42.9'')$$

Inserting (42.8), (42.9'), and (42.9'') into (42.4), we arrive at

$$\mathcal{M} = [4m^2 \hat{a}(q) + 2q^2 \hat{a}(q)] J + i[4ma_\nu(q) + 2i(\gamma_\nu \hat{a}(q) \hat{q} - \hat{q} \hat{a}(q) \gamma_\nu)] J_\nu + 2\gamma_\nu \hat{a}(q) \gamma_\nu J_{\nu\tau}. \quad (42.10)$$

¹⁾ This formula is derived by multiplying the equation

$$\hat{a}(q) \hat{q} = -\hat{q} \hat{a}(q) + 2a(q) q$$

on the left by $\hat{1}$, and using the relation

$$\hat{u}_2 \hat{q} \hat{u}_1 = \hat{u}_2 (\hat{p}_2 - \hat{p}_1) \hat{u}_2 = 0.$$

Let us now calculate the integrals $\underline{J}$, $\underline{J}_0$, $\underline{J}_{0\tau}$. By using the identities¹⁾

$$\frac{1}{ab} = \int_0^1 \frac{dx}{[ax + b(1-x)]^2}, \quad \frac{1}{a^2 b} = \int_0^1 \frac{2x dx}{[ax + b(1-x)]^3}, \quad (42.11)$$

it can be shown that

$$\left. \begin{aligned} \frac{1}{(k^2 - 2p_\mu k)(k^2 - 2p_\nu k)} &= \int_0^1 \frac{dy}{(k^2 - 2p_\mu k)^2} \\ \frac{1}{(k^2 - 2p_\mu k)^2 k^2} &= \int_0^1 \frac{2x dx}{(k^2 - 2p_\mu k)^3} \end{aligned} \right\} \quad (42.12)$$

where

$$p_y = y p_1 + (1-y) p_2, \quad p_\mu = x p_y. \quad (42.12')$$

Therefore, we can write the integrals $\underline{J}$, $\underline{J}_0$, $\underline{J}_{0\tau}$ in the form

$$\left. \begin{aligned} J &= \int_0^1 dy \int_0^1 2x dx \int \frac{d^4 k}{(k^2 - 2p_\mu k)^3}, \\ J_0 &= \int_0^1 dy \int_0^1 2x dx \int \frac{k_\mu d^4 k}{(k^2 - 2p_\mu k)^3}, \\ J_{0\tau} &= \int_0^1 dy \int_0^1 2x dx \int \frac{k_\mu k_\nu d^4 k}{(k^2 - 2p_\mu k)^3}. \end{aligned} \right\} \quad (42.13)$$

The integrations over k in these expressions are performed in the Appendix (see Section 56), where it is shown that

$$\left. \begin{aligned} \int \frac{d^4 k}{(k^2 - 2p_\mu k)^3} &= -\frac{\pi^2 i}{2p_\mu^2}, \\ \int \frac{k_\mu d^4 k}{(k^2 - 2p_\mu k)^3} &= -\frac{\pi^2 i}{2p_\mu^2} p_{\mu\sigma}, \end{aligned} \right\}$$

¹⁾ R. Feynman, Phys. Rev. 76, 769 (1949). In some other cases the identity

$$\frac{1}{a_1 a_2 \dots a_n} = (n-1)! \int_0^1 dx_1 \int_0^{x_1} dx_2 \dots \int_0^{x_{n-2}} dx_{n-1} (a_1 x_{n-1} + a_{n-1} (x_{n-2} - x_{n-1}) + \dots + a_1 (1-x_1))^{-n},$$

is also useful.

$$\int \frac{k_x k_y d^4k}{(k^2 - 2p_x k)^3} = -\frac{\pi^2}{4} \delta_{xy} \left[\ln(-p_x^2) + A_0(N) + \frac{1}{2} \right] - \frac{\pi^2}{2p_x^2} p_{xy} p_{xy} \quad (42.14)$$

The first two of these integrals are absolutely convergent, whereas the third diverges logarithmically for large $|k|$. This last integral is therefore taken over some finite region (N) and contains a constant $A_0(N)$, which behaves as $\ln N$ as $N \rightarrow \infty$ (the number N increases as the region of integration becomes larger).

Using Equation (42.14) and noting that $\underline{p}_x = x \underline{p}_1 + (1-x) \underline{p}_2$, we rewrite (42.13) in the form

$$\left. \begin{aligned} J &= \frac{\pi^2}{2} \int_0^1 I dy, \\ J_x &= \frac{\pi^2}{2} \int_0^1 I_x dy, \\ J_{xy} &= \frac{\pi^2}{2} \int_0^1 I_{xy} dy, \end{aligned} \right\} \quad (42.15)$$

where

$$\left. \begin{aligned} I &= - \int_0^1 \frac{2x dx}{p_x^2}, \\ I_x &= - \int_0^1 \frac{2x p_{xy} dx}{p_x^2} = - \frac{2p_{xy}}{p_y^2}, \\ I_{xy} &= - \int_0^1 2x \left\{ \delta_{xy} \left[\frac{1}{2} A_0(N) + \frac{1}{4} + \frac{1}{2} \ln(-p_x^2) \right] + \frac{p_{xy} p_{xy}}{p_x^2} \right\} dx = \\ &= - \frac{1}{2} \delta_{xy} \left[A_0(N) - \frac{1}{2} + \ln(-p_y^2) \right] - \frac{p_{xy} p_{xy}}{p_y^2}. \end{aligned} \right\} \quad (42.15)$$

We note that I diverges logarithmically for small x . This divergence, however, is not at all connected with the divergences related to high momenta of the virtual particles, but, as will be seen later, is connected with the infrared catastrophe. This, as we know (see Section 32), is due to the inapplicability of perturbation theory to emission and absorption of low-momentum photons.

We shall therefore restrict our considerations in this Section to the interaction of the electron with photons whose frequency is greater than some minimum value $\omega_{\min}$. In practice it is more convenient, however, to perform the integration over photon momenta using another method which is equivalent to the condition $\omega > \omega_{\min}$; this method does not involve restricting the photon frequencies, but assumes that the photon has some extremely small nonzero mass λ . In Section 32 we established the relation between λ and $\omega_{\min}$, and saw that if the electron momentum is small compared to m , then

$$\ln \lambda = \ln 2\omega_{\min} + \frac{5}{6}.$$

In assigning a nonzero mass λ to the photon, we must replace the function $D^F(p)$ as defined by (24.5) by the function

$$D_\lambda^F(p) = -\frac{2i}{(2\pi)^4} \frac{1}{p^2 + \lambda^2}, \quad \lambda^2 > 0.$$

Then instead of expression (42.5) for J we obtain

$$J = \int \frac{d^4k}{(k^2 - 2p_x k)(k^2 - 2p_y k)(k^2 + \lambda^2)}.$$

Noting that

$$\frac{1}{(k^2 - 2p_x k)(k^2 + \lambda^2)} = \int_0^1 \frac{2x dx}{(k^2 - 2p_x x + L_x)^2}, \quad p_x = x p_y, \quad L_x = (1-x)\lambda^2$$

and

$$\int \frac{d^4k}{(k^2 - 2p_x x + L_x)^2} = \frac{\pi^2}{2} \frac{1}{L_x - p_y^2},$$

we write J in the form

$$J = \frac{\pi^2}{2} \int_0^1 I(\lambda) dy,$$

where I is no longer given by (42.15) but by

$$I(\lambda) = \int_0^1 \frac{2x dx}{(1-x)\lambda^2 - x^2 p_y^2},$$

which does not diverge as $x \rightarrow 0$. Performing this integration, we obtain (neglecting terms of order λ)

$$I(\lambda) = -\frac{2}{p_y^2} (\ln \sqrt{-p_y^2} - \ln \lambda)$$

and therefore

$$J = -\frac{\pi^2}{2} \int_0^1 \frac{dy}{p_y^2} (\ln \sqrt{-p_y^2} - \ln \lambda). \quad (42.16)$$

We thus see that the divergence as $y \rightarrow 0$ in expression (42.15) for J is removed by assigning a nonzero rest mass to the photon. It then follows that this divergence is indeed related to the infrared catastrophe, as was stated above.

We shall later eliminate the photon mass λ in (42.16), and consider the interaction of the electron with long-wavelength photons.

Let us transform (42.16) to another form. Noting that $p_1^2 = p_2^2 = -m^2$ and $2p_1 p_2 = -q^2 (q = p_2 - p_1)$, we can write p_y^2 in the form

$$-p_y^2 = -|p_1 + q(1-y)|^2 = m^2 + q^2(1-y)y = m^2[1 - 4y(1-y)\sin^2\theta], \quad (42.17)$$

where θ is related to q by

$$-q^2 = 4m^2\sin^2\theta. \quad (42.17')$$

Eliminating y by means of a new variable ξ

$$2y - 1 = \frac{1}{\tan\theta} \tan\xi,$$

we obtain

$$-p_y^2 = m^2 \frac{\cos^2\theta}{\cos^2\xi}$$

and

$$J = -\frac{2\pi^2 i}{m^2 \sin 2\theta} \left(-\int_0^{\theta} \ln \frac{\cos \xi}{\cos \theta} d\xi - \int_0^{\theta} \ln \frac{\lambda}{m} d\xi \right).$$

Since

$$\int_0^{\theta} \ln \frac{\cos \xi}{\cos \theta} d\xi = \int_0^{\theta} \xi \tan \xi d\xi,$$

we finally arrive at

$$J = -\frac{2\pi^2 i}{m^2 \sin 2\theta} \left(\int_0^{\theta} \xi \tan \xi d\xi + \theta \ln \frac{\lambda}{m} \right). \quad (42.16')$$

Let us now calculate J_0 and J_{0+} as given by (42.13') and (42.15). Noting that

$$p_{y\pm} = y p_{1\pm} + (1-y) p_{2\pm} = \frac{1}{2} (p_{1\pm} + p_{2\pm}) + \frac{\tan \xi}{2 \tan \theta} (p_{1\pm} - p_{2\pm}),$$

we obtain

$$\left. \begin{aligned} J_0 &= \frac{\pi^2 i}{m^2} \frac{0}{\sin 2\theta} (p_{1\pm} + p_{2\pm}), \\ J_{0+} &= \frac{\pi^2 i}{2} \left\{ \delta_{\pm} \left[-\frac{1}{2} A_0(N) + \frac{1}{4} - \ln m \right] + \frac{0 (p_{1\pm} + p_{2\pm}) (p_{1\pm} + p_{2\pm})}{2m^2 \sin 2\theta} + \right. \\ &\quad \left. + \left(\delta_{\pm} + \frac{q d\xi}{q^2} \right) (1 - \theta \cot \theta) \right\}. \end{aligned} \right\} \quad (42.18)$$

Inserting (42.16') and (42.18) into (42.10), we obtain the following expression for $\mathfrak{H}$:

$$\mathfrak{H} = (4m^2 + 2q^2) \hat{a}(q) J + \left[i m \hat{a}(q) (p_1 + p_2) - 2(\hat{p}_1 + \hat{p}_2) \hat{a}(q) \hat{q} + \right. \\ \left. + 2\hat{q} \hat{a}(q) (\hat{p}_1 + \hat{p}_2) + \frac{1}{2} (\hat{p}_1 + \hat{p}_2) \hat{a}(q) (\hat{p}_1 + \hat{p}_2) \right] K_1 + \\ \left. + 2\hat{q} \hat{a}(q) \hat{q} K_2 - 4\hat{a}(q) K_3, \quad (42.19)\right.$$

where

$$\left. \begin{aligned} K_1 &= \frac{\pi^2 i}{m^2} \frac{0}{\sin 2\theta}, \\ K_2 &= -\frac{\pi^2 i}{2\theta^2} (1 - \theta \cot \theta), \\ K_3 &= \frac{\pi^2 i}{2} \left[1 - \theta \cot \theta - \frac{1}{2} A_0(N) + \frac{1}{4} - \ln m \right]. \end{aligned} \right\}$$

Replacing the operator $\hat{p}_1$ on the right and the operator $\hat{p}_2$ on the left by $-m$ as we did in deriving (42.9), we can write the expression in square brackets in (42.19) in the form

$$4im\hat{a}(q)(p_1 + p_2) - 2(\hat{p}_1 + \hat{p}_2)\hat{a}(q)\hat{q} + 2\hat{q}\hat{a}(q)(\hat{p}_1 + \hat{p}_2) + \\ + \frac{1}{2}(\hat{p}_1 + \hat{p}_2)\hat{a}(q)(\hat{p}_1 + \hat{p}_2) = -\frac{1}{2}(20m^2 + 7q^2)\hat{a}(q) + im[\hat{q}\hat{a}(q) - \hat{a}(q)\hat{q}].$$

Inserting this expression into (42.19) and noting that according to (42.17')

$$4m^2 + q^2 = 4m^2 \cos^2 \theta,$$

we finally obtain

$$\mathfrak{H} = 4\pi^2 i \hat{a}(q) \left\{ \frac{2\theta}{\tan 2\theta} \left(\ln \frac{m}{k} - 1 \right) - \frac{2}{\tan 2\theta} \int_0^{\theta} \xi \tan \xi d\xi - \frac{\theta}{2} \tan \theta + \frac{1}{4} A_0(N) - \right. \\ \left. - \frac{3}{8} + \frac{1}{2} \ln m \right\} + \frac{\pi^2}{2m} [\hat{a}(q)\hat{q} - \hat{q}\hat{a}(q)] \frac{2\theta}{\sin 2\theta}. \quad (42.2)$$

We thus see that the divergence as $g \rightarrow 0$ in expression (42.15) for J is removed by assigning a nonzero rest mass to the photon. It then follows that this divergence is indeed related to the Infrared catastrophe, as was stated above.

We shall later eliminate the photon mass λ in (42.16), and consider the interaction of the electron with long-wavelength photons.

Let us transform (42.16) to another form. Noting that $\hat{p}_1^2 = \hat{p}_2^2 = -m^2$ and $2\hat{p}_1 \hat{q} = -\hat{q}_2 (\hat{q} = \hat{p}_2 - \hat{p}_1)$, we can write $\hat{p}_1 \hat{q}$ in the form

$$-\hat{p}_1 \hat{q} = -[\hat{p}_1 + q(1-y)]^2 = m^2 + q^2(1-y)y = m^2[1 - 4y(1-y)\sin^2 \theta], \quad (42.17)$$

where θ is related to $\hat{q}$ by

$$-q^2 = 4m^2 \sin^2 \theta. \quad (42.17)$$

Eliminating y by means of a new variable ξ

$$2y - 1 = \frac{1}{\tan \theta} \tan \xi,$$

we obtain

$$-\hat{p}_1 \hat{q} = m^2 \frac{\cos^2 \theta}{\cos^2 \xi}$$

and

$$J = \frac{2\pi i}{m^2 \sin 2\theta} \left\{ - \int_0^{\frac{\pi}{2}} \ln \frac{\cos \xi}{\cos \theta} d\xi - \int_0^{\frac{\pi}{2}} \ln \frac{\lambda}{m} d\xi \right\}.$$

Since

$$\int_0^{\frac{\pi}{2}} \ln \frac{\cos \xi}{\cos \theta} d\xi = \int_0^{\frac{\pi}{2}} \tan \xi d\xi,$$

we finally arrive at

$$J = -\frac{2\pi i}{m^2 \sin 2\theta} \left\{ \int_0^{\frac{\pi}{2}} \tan \xi d\xi + \theta \ln \frac{\lambda}{m} \right\}. \quad (42.16')$$

Let us now calculate J_{α} and $J_{\alpha\tau}$ as given by (42.13') and (42.15). Noting that

$$p_{1\alpha} = y p_{1\alpha} + (1-y) p_{2\alpha} = \frac{1}{2} (p_{1\alpha} + p_{2\alpha}) + \frac{\tan \xi}{2 \tan \theta} (p_{1\alpha} - p_{2\alpha}),$$

we obtain

$$\begin{aligned} J_{\alpha} &= \frac{\pi^2 i}{m^2 \sin 2\theta} (p_{1\alpha} + p_{2\alpha}), \\ J_{\alpha\tau} &= \frac{\pi^2 i}{2} \left\{ \delta_{\alpha\tau} \left[-\frac{1}{2} A_0(N) + \frac{1}{4} - \ln m \right] + \frac{\theta (p_{1\alpha} + p_{2\alpha}) (p_{1\tau} + p_{2\tau})}{2m^2 \sin 2\theta} + \right. \\ &\quad \left. + \left(\delta_{\alpha\tau} + \frac{q_{\alpha} q_{\tau}}{q^2} \right) (1 - \theta \cot \theta) \right\}. \end{aligned} \quad (42.18)$$

Inserting (42.16') and (42.18) into (42.10), we obtain the following expression for $\mathfrak{H}$:

$$\begin{aligned} \mathfrak{H} &= (4m^2 + 2q^2) \hat{a}(\hat{q}) J + \left[i m \hat{a}(\hat{q}) (\hat{p}_1 + \hat{p}_2) - 2(\hat{p}_1 + \hat{p}_2) \hat{a}(\hat{q}) \hat{q} + \right. \\ &\quad \left. + 2\hat{q} \hat{a}(\hat{q}) (\hat{p}_1 + \hat{p}_2) + \frac{1}{2} (\hat{p}_1 + \hat{p}_2) \hat{a}(\hat{q}) (\hat{p}_1 + \hat{p}_2) \right] K_1 + \\ &\quad + 2\hat{q} \hat{a}(\hat{q}) \hat{q} K_2 - 4\hat{a}(\hat{q}) K_3, \end{aligned} \quad (42.19)$$

where

$$\begin{aligned} K_1 &= \frac{\pi^2 i}{m^2 \sin 2\theta}, \\ K_2 &= -\frac{\pi^2 i}{2q^2} (1 - \theta \cot \theta), \\ K_3 &= \frac{\pi^2 i}{2} \left[1 - \theta \cot \theta - \frac{1}{2} A_0(N) + \frac{1}{4} - \ln m \right]. \end{aligned}$$

Replacing the operator $\hat{p}_1$ on the right and the operator $\hat{p}_2$ on the left by $-m$ as we did in deriving (42.9), we can write the expression in square brackets in (42.19) in the form

$$\begin{aligned} &4ima(\hat{q}) (\hat{p}_1 + \hat{p}_2) - 2(\hat{p}_1 + \hat{p}_2) \hat{a}(\hat{q}) \hat{q} + 2\hat{q} \hat{a}(\hat{q}) (\hat{p}_1 + \hat{p}_2) + \\ &+ \frac{1}{2} (\hat{p}_1 + \hat{p}_2) \hat{a}(\hat{q}) (\hat{p}_1 + \hat{p}_2) = -\frac{1}{2} (20m^2 + 7q^2) \hat{a}(\hat{q}) + im [\hat{q} \hat{a}(\hat{q}) - \hat{a}(\hat{q}) \hat{q}]. \end{aligned}$$

Inserting this expression into (42.19) and noting that according to (42.17)

$$4m^2 + q^2 = 4m^2 \cos^2 \theta,$$

we finally obtain

$$\begin{aligned} \mathfrak{H} &= 4\pi^2 \hat{a}(\hat{q}) \left\{ \frac{2\theta}{\sin 2\theta} \left(\ln \frac{m}{\lambda} - 1 \right) - \frac{2}{\tan 2\theta} \int_0^{\frac{\pi}{2}} \tan \xi d\xi - \frac{\theta}{2} \tan \theta + \frac{1}{4} A_0(N) - \right. \\ &\quad \left. - \frac{3}{8} + \frac{1}{2} \ln m \right\} + \frac{\pi^2}{2m} [\hat{a}(\hat{q}) \hat{q} - \hat{q} \hat{a}(\hat{q})] \frac{2\theta}{\sin 2\theta}. \end{aligned} \quad (42.2)$$

This expression must now be renormalized. Since $\mathcal{U}$ diverges logarithmically, $\mathcal{U}$ is regularized by subtracting its value at $q = 0$, that is at $q = 0$, from the factor multiplying $\underline{a}(q)$, as was shown in Section 26 [see (26.11); in our case p_1 and p_2 are free-electron momenta].

The regularized expression for $\mathcal{U}$ is

$$\mathcal{U}_R = 4\pi^2 i \underline{a}(q) \left[\left(\frac{20}{\sqrt{\tan 2\theta}} - 1 \right) \left(\ln \frac{m}{\lambda} - 1 \right) - \frac{2}{\tan 2\theta} \int_0^{\frac{\pi}{2}} \xi \tan \xi d\xi - \frac{0}{2} \tan \theta \right] + \frac{\pi^2}{2m} [\hat{a}(q) \hat{q} - \hat{q} \hat{a}(q)] \frac{20}{\sin 2\theta}. \quad (42.21)$$

Inserting (42.21) into (42.3), we obtain the matrix element corresponding to Fig. 53, diagram 1, namely

$$M_1^{(9)} = i \frac{e^2}{(2\pi)^4} \bar{u}_2 \mathcal{U}_R u_1, \quad (42.21')$$

which does not take into account the interaction between electrons and long-wavelength photons.

Before analyzing this expression, we shall calculate $M_2^{(9)}$.

§ 43. Vacuum Polarization

1. Calculation of the Vacuum Polarization Matrix Element

The matrix element $M_2^{(9)}$ which describes the polarization of the electron-positron vacuum (diagram 2 on Fig. 53) is given by Equation (42.2). Introducing the notation

$$T_{\mu\nu} = \int \text{Sp} \left\{ \frac{i(\hat{p} + \hat{q}) - m}{(p+q)^2 + m^2} \gamma_\mu \frac{i\hat{p} - m}{p^2 + m^2} \gamma_\nu \right\} d^4 p, \quad (43.1)$$

we shall rewrite $M_2^{(9)}$ in the form

$$M_2^{(9)} = -i \frac{e^2}{(2\pi)^4} \bar{u}_2 \left(\gamma_\mu \frac{\underline{a}(q)}{q^2} \gamma_\nu \right) T_{\mu\nu} u_1. \quad (43.2)$$

The tensor $T_{\mu\nu}$ has a simple physical meaning. Indeed, replacing Fig. 53, diagram 2, by a skeleton diagram which does not contain the photon self-energy part, we must replace the external potential $\underline{a}_\mu^{(9)}(x)$ by a new potential $\delta \underline{a}_\mu(x)$ which will give the matrix element $M_2^{(9)}$ to first order:

$$M_2^{(9)} = e \bar{u}_2 \gamma_\mu \delta a_\mu(q) u_1, \quad \delta a_\mu(q) = \int \delta A_\mu(x) e^{-iqx} d^4 x. \quad (43.3)$$

⁵⁾ R. Feynman. See the Reference on p. 447.

Comparison of this equation with (43.2) shows that

$$\delta a_\mu(q) = -i \frac{e^2}{(2\pi)^4} \frac{1}{q^2} T_{\mu\nu} \gamma_\nu \underline{a}(q). \quad (43.4)$$

The potential $\delta \underline{A}_\mu(x)$ should be considered a correction, due to the polarization of the electron-positron vacuum, to the "given" external potential $\underline{a}_\mu^{(9)}(x)$. If we apply the operator $-\square$ to $\delta \underline{A}_\mu(x)$, we obtain the correction to the external current $\underline{j}_\mu^{(9)}(x)$ which is due to its interaction with the zero-point oscillations of the electron-positron field. Denoting the Fourier transform of this correction by $\delta \underline{j}_\mu(q)$, Equation (43.4) leads to

$$\delta j_\mu(q) = -i \frac{e^2}{(2\pi)^4} T_{\mu\nu} \gamma_\nu \underline{a}(q). \quad (43.5)$$

Let us now calculate $T_{\mu\nu}$, the tensor which determines vacuum polarization. Replacing p by $p - \frac{1}{2} q$ in (43.1), we write $T_{\mu\nu}$ in the form

$$\begin{aligned} T_{\mu\nu} &= \int \text{Sp} \left\{ \frac{i(\hat{p} + \frac{1}{2}\hat{q}) - m}{(p + \frac{1}{2}q)^2 + m^2} \gamma_\mu \frac{i(\hat{p} - \frac{1}{2}\hat{q}) - m}{(p - \frac{1}{2}q)^2 + m^2} \gamma_\nu \right\} d^4 p = \\ &= K \text{Sp} \left\{ \left(i\frac{\hat{q}}{2} - m \right) \gamma_\mu \left(-i\frac{\hat{q}}{2} - m \right) \gamma_\nu \right\} + \\ &+ iK_s \text{Sp} \left\{ \left(i\frac{\hat{q}}{2} - m \right) \gamma_\mu \gamma_5 \gamma_\nu \left(-i\frac{\hat{q}}{2} - m \right) \gamma_5 \right\} - K_{st} \text{Sp} \left\{ \gamma_\mu \gamma_5 \gamma_\nu \gamma_5 \right\}, \quad (43.6) \end{aligned}$$

where

$$\begin{aligned} K &= \int \frac{d^4 p}{\left[\left(p + \frac{1}{2}q \right)^2 + m^2 \right] \left[\left(p - \frac{1}{2}q \right)^2 + m^2 \right]}, \\ K_s &= \int \frac{p_\alpha d^4 p}{\left[\left(p + \frac{1}{2}q \right)^2 + m^2 \right] \left[\left(p - \frac{1}{2}q \right)^2 + m^2 \right]}, \\ K_{st} &= \int \frac{p_\alpha p_\beta d^4 p}{\left[\left(p + \frac{1}{2}q \right)^2 + m^2 \right] \left[\left(p - \frac{1}{2}q \right)^2 + m^2 \right]}. \quad (43.7) \end{aligned}$$

To calculate these integrals we set $\underline{x} = \frac{1}{2}(1 + \eta)$ in the first of Equations (42.11), obtaining

$$\frac{1}{ab} = \frac{1}{2} \int_{-1}^{+1} \frac{d\eta}{\left(\frac{1+\eta}{2} + \frac{1-\eta}{2} b \right)^2}.$$

It then follows that

$$\frac{1}{\left[\left(p+\frac{1}{2}q\right)^2+m^2\right]\left[\left(p-\frac{1}{2}q\right)^2+m^2\right]} = \frac{1}{2} \int_{-1}^{+1} \frac{d\eta}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2},$$

where

$$L = \frac{1}{4}q^2(1-\eta^2) + m^2.$$

Therefore

$$\left. \begin{aligned} K &= \frac{1}{2} \int_{-1}^{+1} d\eta \int \frac{d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2}, \\ K_s &= \frac{1}{2} \int_{-1}^{+1} d\eta \int \frac{p_s d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2}, \\ K_{st} &= \frac{1}{2} \int_{-1}^{+1} d\eta \int \frac{p_s p_t d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2}. \end{aligned} \right\} \quad (43.7)$$

The integrals over p in these expressions are calculated in the Appendix (see Section 56) where it is shown that

$$\int \frac{d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2} = -\pi^2 i \left[\ln L + A_0(N) \right], \quad (43.8)$$

$$\int \frac{p_s d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2} = -\pi^2 i \frac{1}{2} q_s \eta \left[\ln L + A_0(N) + \frac{1}{2} \right], \quad (43.8')$$

$$\int \frac{p_s p_t d^4p}{\left[\left(p-\frac{1}{2}q\eta\right)^2+L\right]^2} = \frac{\pi^2 i}{4} \delta_{st} \left\{ L \ln L - L + \left(L + \frac{1}{4} q^2 \eta^2 \right) \left[A_0(N) + \frac{1}{2} \right] - \frac{1}{4} q^2 \eta^2 \left[A_0(N) + \frac{5}{6} \right] + \frac{1}{2} A_2(N) \right\} - \frac{\pi^2 i}{4} q_s q_t \eta^2 \left[\ln L + A_0(N) + \frac{5}{6} \right]. \quad (43.8'')$$

All these integrals diverge for large $|p|$, and therefore the integral is taken over some large but finite region N . The constants A_0 and A_2 in (43.8) behave as $\ln N$ and N^2 for large N .

Equations (43.8), (43.8'), and (43.8'') must now be averaged over η (with $-1 \leq \eta \leq 1$) and inserted into (43.6), after which the tensor $T_{\mu\nu}$ must be renormalized. This renormalization is performed by subtracting the first three terms in the expansion of $T_{\mu\nu}$ in a power series in q . We note that this subtraction is in agreement

with the fact that according to (43.2) $M_2^{(4)}$ contains q^2 in the denominator; since $M_2^{(4)}$ must be finite for $q=0$, the tensor $T_{\mu\nu}$ should have at least a third-order zero at $q=0$.

This subtraction is also in agreement with the concept of charge renormalization. Indeed, subtraction of a term proportional to q^2 from $T_{\mu\nu}$ according to (43.5), is the same as subtracting from δJ_μ a term proportional to the original "external" current $J_\mu^{(0)}$ (see Section 27).

Since K , K_s , K_{st} are multiplied in (43.6) by expressions which are, respectively, quadratic in q , linear in q , and independent of q , the term $-\pi^2 i A_0(N)$ can be dropped in (43.8), the term $-\pi^2 i \frac{1}{2} q_s q_t \eta^2 \left[A_0(N) + \frac{5}{6} \right]$ can be dropped in (43.8''), and the term

$$\frac{\pi^2 i}{2} \delta_{st} \left\{ -L + \left(L + \frac{1}{4} q^2 \eta^2 \right) \left[A_0(N) + \frac{1}{2} \right] - \frac{1}{4} q^2 \eta^2 \left[A_0(N) + \frac{5}{6} \right] + \frac{1}{2} A_2(N) \right\} - \frac{\pi^2 i}{4} q_s q_t \eta^2 \left[A_0(N) + \frac{5}{6} \right]$$

can be dropped in (43.8'').

After these subtractions we obtain the following expressions for K , K_s , K_{st} :

$$\left. \begin{aligned} K &= -\frac{\pi^2 i}{2} \int_{-1}^{+1} \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta, \\ K_s &= -\frac{\pi^2 i}{4} q_s \int_{-1}^{+1} \eta \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta, \\ K_{st} &= -\frac{\pi^2 i}{8} q_s q_t \int_{-1}^{+1} \eta^2 \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta + \\ &\quad + \frac{\pi^2 i}{4} \delta_{st} \int_{-1}^{+1} \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta. \end{aligned} \right\} \quad (43.7'')$$

Since the integrand of K_s is an odd function of η , we obtain, as may have been expected (compare (26.10),

$$K_s = 0.$$

Noting further that

$$\int_{-1}^{+1} \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta = 4 \ln m - 4 (1 - \theta \cot \theta) = R_1,$$

$$\int_{-1}^{+1} \eta^2 \ln \left[\frac{1}{4} (1-\eta^2) q^2 + m^2 \right] d\eta = \frac{4}{3} \ln m - \frac{4}{9} + \frac{4}{3} \cot^2 \theta (1 - \theta \cot \theta) = R_2,$$

where θ is related to q^2 by $-q^2 = 4m^2 \sin^2 \theta$, we can write K and $K_{\mu\nu}$ in the form

$$K = -\frac{1}{2} \pi^2 i R_1, \\ K_{\mu\nu} = -\frac{1}{8} \pi^2 i q_\mu q_\nu R_2 + \frac{1}{4} \pi^2 i g_{\mu\nu} \left[\left(m^2 + \frac{1}{4} q^2 \right) R_1 - \frac{1}{4} q^2 R_2 \right].$$

According to the above concepts we can drop terms independent of θ ($q = 0$) in R_1 and R_2 , since they vanish when $T_{\mu\nu}$ is renormalized. Making use of Equation (3.23) we can write $T_{\mu\nu}$ in the form

$$T_{\mu\nu} = 4\pi^2 i (q_\mu q_\nu - \delta_{\mu\nu} q^2) \left(\frac{4m^2 - 2q^2}{-3q^2} \right) (1 - \theta \cot \theta).$$

2. Renormalization of the Matrix Element.

We must now renormalize $T_{\mu\nu}$; that is, we must subtract from it the first three terms of its power series expansion in q_0 .

Since $T_{\mu\nu}$ contains the factor $g_\mu q_\nu - \delta_{\mu\nu} q^2$, which is quadratic in q_0 , renormalization reduces to subtracting a constant from the factor $\frac{4m^2 - 2q^2}{-3q^2} (1 - \theta \cot \theta)$. Noting that $\theta = \arctan \left(\frac{-q^2}{4m^2 + q^2} \right)^{1/2}$, it is easy to see that this constant is $\frac{1}{9}$. Therefore, the final renormalized expression for $T_{\mu\nu}$ is ¹⁾

$$T_{\mu\nu R} = 4\pi^2 i (q_\mu q_\nu - \delta_{\mu\nu} q^2) \left[\frac{4m^2 - 2q^2}{-3q^2} (1 - \theta \cot \theta) - \frac{1}{9} \right]. \quad (43.9)$$

If we multiply $T_{\mu\nu R}$ by $-i \frac{e^2}{(2\pi)^4} \frac{1}{q^2} a_\nu(q)$, then according to (43.5) we obtain the correction $J_{\mu}^{(p)}(q)$ to the Fourier transform of the original "external" current $J_{\mu}^{(0)}(q)$ which is due to its interaction with the zero-point oscillations of the electron-positron field:

$$J_{\mu}^{(p)}(q) = \frac{1}{(2\pi)^4} e^2 (q_\mu q_\nu - \delta_{\mu\nu} q^2) \left[\frac{4m^2 - 2q^2}{-3q^2} (1 - \theta \cot \theta) - \frac{1}{9} \right] a_\nu(q). \quad (43.10)$$

It is easily shown that this correction satisfies conservation of charge

$$q_\mu J_{\mu}^{(p)}(q) = 0.$$

¹⁾ R. Feynman. See the reference on p. 447.

Noting that $q_\nu a_\nu(q) = 0$ and $q^2 a_\mu(q) = J_{\mu}^{(0)}(q)$, we rewrite $J_{\mu}^{(p)}(q)$ in the form

$$J_{\mu}^{(p)}(q) = -\frac{e^2}{(2\pi)^4} \left[\frac{4m^2 - 2q^2}{-3q^2} (1 - \theta \cot \theta) - \frac{1}{9} \right] J_{\mu}^{(0)}(q). \quad (43.11)$$

Expanding $T_{\mu\nu R}$ in a power series in q_0 and considering only the fourth-order terms, we obtain

$$T_{\mu\nu R} = -\frac{4\pi^2 i}{15m^2} (q_\mu q_\nu - \delta_{\mu\nu} q^2) q^2. \quad (43.12)$$

In this case the correction to the Fourier transform of the current is

$$J_{\mu}^{(p)}(q) = -\frac{e^2}{60\pi^2 m^2} q^2 J_{\mu}^{(0)}(q). \quad (43.13)$$

From this we obtain the following expression for the current as a function of the space-time variables:

$$J_{\mu}^{(p)}(x) = -\frac{e^2}{60\pi^2 m^2} \square J_{\mu}^{(0)}(x). \quad (43.14)$$

If we keep sixth-order terms in the expansion of $T_{\mu\nu R}$ in a power series in q_0 then instead of (43.12) we obtain

$$J_{\mu}^{(p)}(x) = -\frac{e^2}{60\pi^2 m^2} \square J_{\mu}^{(0)}(x) - \frac{e^2}{680\pi^2} \left(\frac{1}{m^2} \square \right)^2 J_{\mu}^{(0)}(x).$$

Inserting Equation (43.9) for $T_{\mu\nu R}$ into (43.2), we obtain the following expression for the matrix element $M_{\mu}^{(0)}$:

$$M_{\mu}^{(0)} = -i \frac{e^2}{(2\pi)^4} a_0 \left(\gamma_\mu \frac{a_\nu(q)}{q^2} T_{\mu\nu R} \right) u_1. \quad (43.15)$$

To determine the probability for the processes represented by the diagrams of Fig. 53, we clearly need to find the total matrix element $S_{\mu}^{(0)} = M_{\mu}^{(0)} + M_{\mu}^{(p)}$. Using Equations (42.21), (42.21), (43.13), and (43.9), we can write the total matrix element in the form

$$S_{\mu}^{(0)} = M_{\mu}^{(0)} + M_{\mu}^{(p)} = u_2 S^{(0)}(q) u_1, \quad (43.16)$$

where

$$S^{(0)}(q) = i \frac{e^2}{(2\pi)^4} \left[\hat{u}_R - \gamma_\mu \frac{a_\mu(q)}{q^2} T_{\mu,R} \right] = \\ = - \frac{e^2}{(2\pi)^4} \left\{ \hat{a}(q) \left[\frac{2\theta}{\sin 2\theta} - 1 \right] \left(\ln \frac{m}{\lambda} - 1 \right) - \frac{\theta}{2} \tan \theta - \frac{2}{\tan 2\theta} \int_0^\theta \xi \tan \xi d\xi \right\} + \\ + \frac{i}{8m} \left[\hat{q} \hat{a}(q) - \hat{a}(q) \hat{q} \right] \frac{2\theta}{\sin 2\theta} - \frac{\gamma_\mu a_\mu(q)}{q^2} (q_\mu \gamma_\mu - \delta_{\mu\mu} q^2) \times \\ \times \left[\frac{4m^2 - 2q^2}{3q^2} (1 - \theta \cot \theta) - \frac{1}{9} \right]. \quad (43.15)$$

Later we shall need an expansion of $S^{(0)}(q)$ in powers of q_σ up to quadratic terms. Such an expansion for $\mathcal{U}_k$ is given by

$$\mathcal{U}_R = \frac{4\pi i}{3m^2} q^2 \hat{a}(q) \left(\ln \frac{m}{\lambda} - \frac{3}{8} \right) - \frac{\pi^2}{2m} [\hat{q} \hat{a}(q) - \hat{a}(q) \hat{q}]. \quad (43.16)$$

Inserting (43.16) and (43.9) into (43.15), we obtain up to third-order terms

$$S^{(0)}(q) = - \frac{e^2}{(4\pi)^3} \left\{ \frac{4}{3m^2} q^2 \hat{a}(q) \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) + \frac{i}{2m} [\hat{q} \hat{a}(q) - \hat{a}(q) \hat{q}] \right\} \quad (43.17)$$

(we have made use of the fact that $\hat{q}_\nu \hat{a}_\nu(q) = 0$).

§ 4.4. Effective Electron Potential Energy. Magnetic Moment of the Electron.

1. Effective Electron Potential Energy.

In the preceding Section we have calculated the matrix element $S_{1 \rightarrow 2}^{(0)}(q)$, which gives the radiative corrections to electron scattering in an external field $A_\mu^{(0)}(x)$. If we add to $S_{1 \rightarrow 2}^{(0)}(q)$ the matrix element $S_{1 \rightarrow 2}^{(1)}(q) = \bar{u}_2 e \hat{A}_\mu(q) u_1 (q = p_2 - p_1)$, which gives electron scattering in the first approximation, the matrix element $S_{1 \rightarrow 2}^{(1)}(q)$ so obtained gives the interaction of the electron with the zero-point oscillations of the field up to terms of order e^4 :

$$S_{1 \rightarrow 2}^{(1)}(q) = \bar{u}_2 e \hat{a}(q) u_1 + \bar{u}_2 S^{(0)}(q) u_1 = \bar{u}_2 \{ \gamma e a(q) + i \not{q} [\gamma \varphi(q) + \delta U(q)] \} u_1, \quad (44.1)$$

where

$$\delta U(q) = \frac{1}{i} \beta S^{(0)}(q). \quad (44.2)$$

Equation (44.1) shows that $\delta U(q)$ may be treated as the Fourier transform of the correction to the electron potential energy due to interaction between the electron and the zero-point field oscillations. We shall call this correction

the effective potential energy, which gives the interaction of the electron with the zero-point field oscillations.

Since $S^{(0)}(q)$, according to (43.15), can be written

$$S^{(0)}(q) = F_+(q) a_+(q),$$

where

$$F_+(q) = - \frac{e^2}{(2\pi)^4} \left\{ \gamma_\mu \left[\frac{2\theta}{\sin 2\theta} - 1 \right] \left(\ln \frac{m}{\lambda} - 1 \right) - \frac{\theta}{2} \tan \theta - \right. \\ \left. - \frac{2}{\tan 2\theta} \int_0^\theta \xi \tan \xi d\xi \right\} + \frac{i}{8m} \hat{q} \gamma_\mu - \gamma_\mu \hat{q} \frac{2\theta}{\sin 2\theta} - \gamma_\mu \frac{1}{q^2} (q_\mu \gamma_\mu - \\ - \delta_{\mu\mu} q^2) \left[\frac{4m^2 - 2q^2}{3q^2} (1 - \theta \cot \theta) - \frac{1}{9} \right], \quad (44.3)$$

we may write

$$\delta U(q) = \frac{1}{i} \beta F_+(q) a_+(q). \quad (44.4)$$

If we know its Fourier transform $\delta U(x)$, we can find the effective electron potential energy $\delta U(x)$ as a function of the space-time coordinates, namely

$$\delta U(x) = \frac{1}{(2\pi)^4 i} \int \beta F_+(q) a_+(q) e^{iqx} d^4q. \quad (44.5)$$

Noting that

$$a_+(q) = \int A_\nu^{(0)}(y) e^{-iqy} d^4y,$$

we can write (44.5) in the form

$$\delta U(x) = \frac{1}{(2\pi)^4 i} \int \beta F_+(q) e^{iq(x-y)} A_\nu^{(0)}(y) d^4q d^4y = \\ = \frac{1}{i} \int \beta F_+(x-y) A_\nu^{(0)}(y) d^4y, \quad (44.6)$$

where

$$F_+(x) = \frac{1}{(2\pi)^4} \int F_+(q) e^{iqx} d^4q.$$

We see that the effective electron potential energy can be written in the form of an integral operator acting on the given external field $A_{\nu}^{(0)}(x)$.

For actual calculations, another form of $\delta U(x)$ is more convenient. Noting that $\frac{q}{\nu} e^{iqx} = \frac{1}{i} \frac{\partial}{\partial x_{\nu}} e^{iqx}$ we can rewrite (44.5) in the form

$$\delta U(x) = \frac{1}{i} \beta F_{\nu} \left(\frac{1}{i} \frac{\partial}{\partial x_{\nu}} \right) \frac{1}{(2\pi)^4} \int d^4q A_{\nu}^{(0)}(x) F_{\nu} \left(\frac{1}{i} \frac{\partial}{\partial x_{\nu}} \right) A_{\nu}^{(0)}(x). \quad (44.7)$$

This equation can be used if $F_{\nu}(q)$ can be written as a convergent power series in q ; then $F_{\nu} \left(\frac{1}{i} \frac{\partial}{\partial x_{\nu}} \right)$ is a differential operator acting on $A_{\nu}^{(0)}(x)$.

It is possible to represent $F_{\nu}(q)$ as a power series in q if the field $A_{\nu}^{(0)}(x)$ does not vary too rapidly in space and time. We shall assume that the external field varies so slowly that we may neglect all terms higher than the quadratic in the expansion of $F_{\nu}(q)$ in a power series in q . According to (43.17) this expression for $F_{\nu}(q)$ can be written

$$F_{\nu}(q) = \frac{e^2}{(4\pi)^2} \left\{ \frac{4}{3m^2} q^2 \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) \gamma_{\nu} + \frac{i}{2m} (\hat{q} \gamma_{\nu} - \gamma_{\nu} \hat{q}) \right\}.$$

Thus the effective potential energy $\delta U(x)$ is

$$\delta U(x) = \frac{e^2}{(4\pi)^2} \left\{ \frac{4}{3m^2} \frac{1}{i} \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) \beta \gamma_{\nu} \square A_{\nu}^{(0)}(x) - \frac{1}{2im} \beta \gamma_{\nu} \gamma_{\mu} F_{\mu\nu}(x) \right\}, \quad (44.8)$$

where $F_{\mu\nu}(x) = \frac{\partial}{\partial x_{\mu}} A_{\nu}^{(0)}(x) - \frac{\partial}{\partial x_{\nu}} A_{\mu}^{(0)}(x)$ is the field tensor.

Let us now introduce the matrices

$$\frac{1}{i} \gamma_1 \gamma_3 = \sigma_3, \quad \frac{1}{i} \gamma_3 \gamma_3 = \sigma_1, \quad \frac{1}{i} \gamma_3 \gamma_1 = \sigma_2,$$

It can easily be shown that

$$\sigma_1 \sigma_2 = i\sigma_3, \quad \sigma_2 \sigma_3 = i\sigma_1, \quad \sigma_3 \sigma_1 = i\sigma_2,$$

or in other words, that the σ_i matrices satisfy the same commutation relations as the electron spin operators. With the aid of these matrices we can write $\delta U(x)$ in the form

$$\delta U(x) = -\frac{e^2}{(4\pi)^2 m} (\beta \sigma H - i \beta \sigma E) + \frac{e^2}{(4\pi)^2 3m^2} \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) (\square \varphi - \square A), \quad (44.9)$$

where E and H are the electric and magnetic fields, and φ and A are the scalar and vector potentials.

2. Radiative Corrections to the Electron Magnetic Moment.

It follows from (44.9) that an electron in a constant magnetic field gains an additional energy

$$-\frac{e^2}{8\pi^2 2m} \beta \sigma H = -\frac{1}{2\pi} \frac{1}{137} \frac{e}{2m} \beta \sigma H, \quad \text{due to its interaction with the zero-point oscillations of the electromagnetic field. But the energy of an electron in a magnetic field is}^{1)}$$

$$U_m = -\mu \beta \sigma H,$$

where μ is the magnetic moment of the electron. We may therefore say that due to the electron's interaction with the zero-point oscillations of the electromagnetic field, it gains an additional magnetic moment equal to $\Delta\mu = \frac{e^2}{8\pi^2} \frac{e}{2m}$, or²⁾

$$\Delta\mu = \frac{1}{2\pi} \frac{e^2}{4\pi\hbar c} \frac{e\hbar}{2mc} = \frac{\alpha}{2\pi} \mu_0, \quad (44.10)$$

where $\mu_0 = \frac{e\hbar}{2mc}$ is the Bohr magneton, and $\alpha = \frac{e^2}{4\pi\hbar c} = \frac{1}{137}$ is the fine structure constant. Thus up to terms of order α^2 , the magnetic moment of the electron is

$$\mu = \mu_0 \left(1 + \frac{\alpha}{2\pi} \right).$$

The "anomalous" magnetic moment $\Delta\mu$ of the electron has been experimentally observed³⁾.

We note that the interaction of the electron with the zero-point field oscillations taking account of fourth-order effects, gives the following value for the magnetic moment of the electron:⁴⁾

$$\mu = \mu_0 \left(1 + \frac{\alpha}{2\pi} - 2.97 \frac{\alpha^2}{\pi^2} \right).$$

Equation (44.9) for the effective potential energy contains the nonexistent photon "mass" λ , which we introduced in connection with the low-frequency divergence of the matrix element $\underline{S}_1^{(0)}$. We shall now show how to remove this "mass" when actually calculating radiative corrections to electron scattering in an external field or radiative shifts of atomic levels.

§ 45. Radiative Corrections to Scattering

1. Radiative Corrections to Scattering of an Electron by an External Field

In Section 13 we found the cross section for electron scattering by an external field. In this section we shall show how to obtain the radiative corrections to the scattering, which are due to the interaction of the electron with the zero-point oscillations of the electromagnetic field and to the polarization of the electron-

¹⁾ In the nonrelativistic approximation β becomes the unit matrix.

²⁾ This result is due to Schwinger [Phys. Rev. 75, 1912 (1949)].

³⁾ See the article by Kusch and Foley in Atomic Electron Level Shifts, p. 52.

⁴⁾ R. Karplus and N. Kroll, Phys. Rev. 77, 536 (1950).

positron vacuum¹⁾. We shall assume that the external field can be treated as a perturbation.

Since radiative corrections are a third-order effect, and since scattering takes place even in the first order, in addition to radiative corrections we must take into account second-order scattering processes (scattering in the second Born approximation).

Diagrams representing the processes of interest are shown in Fig. 54 (diagrams 1 and 2 illustrate the first and second Born approximation, and diagrams 3 and 4 illustrate radiative corrections and vacuum polarization).

The matrix elements corresponding to these diagrams, according to the rules of Section 24, are

$$\left. \begin{aligned} S_{fi}^{(0)} &= e \bar{u}_2 \hat{a}(q) u_1, \\ M_{fi}^{(1)} &= -i \frac{e^2}{(2\pi)^4} \bar{u}_2 \left(\int \hat{a}(p_2 - p) \frac{i \hat{p} - m}{p^2 + m^2} \hat{a}(p - p_1) d^4 p \right) u_1, \\ M_{fi}^{(2)} &= \frac{ie^3}{(2\pi)^4} \bar{u}_2 \left(\int \gamma_k \frac{i(\hat{p}_2 - k) - m}{(p_2 - k)^2 + m^2} \hat{a}(q) \frac{i(\hat{p}_1 - k) - m}{(p_1 - k)^2 + m^2} \gamma_k \frac{d^4 k}{k^2} \right) u_1, \\ M_{fi}^{(3)} &= -\frac{ie^3}{(2\pi)^4} \bar{u}_2 \left(\alpha_\mu(q) \frac{1}{q^2} \int \text{Sp} \left(\frac{i(\hat{p} + \hat{q}) - m}{(p + q)^2 + m^2} \gamma_\mu \frac{i \hat{p} - m}{p^2 + m^2} \right) d^4 p \right) u_1, \end{aligned} \right\} \quad (45.1)$$

where u_1 and u_2 are the spinor amplitudes of the electron states with momenta p_1 and p_2 and $q = p_2 - p_1$.

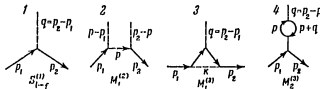


Fig. 54

We shall also be interested in elastic²⁾ scattering of electrons in the Coulomb field of the nucleus. In this case, as we know,

$$\hat{a}(q) = \hat{a}_q 2\pi \delta(\epsilon_1 - \epsilon_2) = 2\pi \beta \frac{Ze}{q^2} \delta(\epsilon_1 - \epsilon_2), \quad (45.2)$$

where Ze is the charge of the nucleus, $q = p_2 - p_1$ and ϵ_1 and ϵ_2 are the electron energies before and after scattering.

Since $\epsilon_1 = \epsilon_2$ and $q_0 = 0$, we may write

$$q^2 = q^2 = 4p^2 \sin^2 \frac{\theta}{2},$$

where θ is the scattering angle, and $|p| = |p_1| = |p_2|$. As in Section 32, we shall make use of the quantity Φ , which is related to q by Equation (32.51), namely

¹⁾ J. Schwinger, Phys. Rev. 75, 1912 (1949).

²⁾ Actually, we shall be interested in scattering which is almost elastic; see below.

$$q^2 = -4m^2 \sinh^2 \Phi, \sinh \Phi = \frac{v}{\sqrt{1-v^2}} \sinh \frac{\theta}{2},$$

where v is the ratio between the electron velocity and the velocity of light. We note that the quantity θ introduced in (42.17) is given by $\theta = i\Phi$ (in the present case it is Φ , rather than θ , which is real).

Inserting (45.2) into the first of Equations (45.1), we obtain the following expression for $S_{fi}^{(1)}$:

$$S_{fi}^{(1)} = 2\pi i \frac{Ze^2}{q^2} (\bar{u}_2 \hat{a} u_1) \delta(\epsilon_2 - \epsilon_1). \quad (45.3)$$

We have calculated $M_{fi}^{(1)}$ and $M_{fi}^{(2)}$ in general form in Sections 42, 43 [see (42.21), (43.2), (43.9)]. In order to find $M_{fi}^{(3)}$ and $M_{fi}^{(4)}$ for the case of Coulomb scattering $\hat{a}$ in Equations (42.21) and (43.2) should be replaced by expression (45.2) and θ should be replaced by $i\Phi$. Noting also that $\hat{a}(q)$ is proportional to B , and that therefore

$$\hat{a}(q) \hat{q} = \hat{q} \hat{a}(q) = 2\hat{a} \hat{q} = 4\pi i \beta \frac{Ze}{q^2} \hat{q} \delta(\epsilon_2 - \epsilon_1),$$

we finally arrive at the following expressions for $M_{fi}^{(3)}$ and $M_{fi}^{(4)}$:

$$\left. \begin{aligned} M_{fi}^{(3)} &= -i \frac{e^3 Ze}{8\pi q^2} \bar{u}_2 \left\{ 4 \left[\left(1 - 2\Phi \coth \Phi \right) \left(1 - \frac{1}{m} \right) \cdot \frac{\Phi}{2} \tanh \Phi + \right. \right. \\ &\quad \left. \left. + 2 \coth 2\Phi \int_0^\Phi u \tanh u du \right] - \frac{1}{m} \frac{2\Phi}{\sinh 2\Phi} \right\} u_1 \delta(\epsilon_1 - \epsilon_2), \\ M_{fi}^{(4)} &= -i \frac{e^3 Ze}{8\pi q^2} \bar{u}_2 \mu_1 \left\{ \left(1 - \Phi \coth \Phi \right) \left(1 - \frac{1}{3} \coth \Phi \right) - \frac{1}{3} \right\} \delta(\epsilon_1 - \epsilon_2). \end{aligned} \right\} \quad (45.4)$$

The matrix element $M_{fi}^{(3)}$ can be determined by the second of equations (45.1). We shall not, however, go into a detailed calculation of the integral in the expression for $M_{fi}^{(3)}$, but shall merely present the final result¹⁾. If $\hat{A}^{(3)}(x)$ is a screened Coulomb potential given by

$$\hat{A}^{(3)}(x) = i\beta \frac{Ze}{4\pi r} e^{-\eta r},$$

where η is the screening constant, then for sufficiently small η

$$M_{fi}^{(3)} = -4Z^2 \alpha^2 \bar{u}_2 \left\{ m(I - J) + \beta \epsilon_1(I + J) \right\} u_1 \delta(\epsilon_1 - \epsilon_2), \quad (45.5)$$

¹⁾ R. Dalitz, Proc. Roy. Soc. A 206, 509 (1951).

where

$$I = \frac{\pi^2 l}{2|p|^3 \sin^2 \frac{\theta}{2}} \ln \left(\frac{2|p| \sin \frac{\theta}{2}}{\eta} \right), \quad (45.5)$$

$$I - J = \frac{\pi^2}{4|p|^3 \cos^2 \frac{\theta}{2}} \left(\frac{1}{\sin \frac{\theta}{2}} - 1 \right) + \frac{\pi^2 l}{4|p|^3 \cos^2 \frac{\theta}{2}} \ln \left(\sin^2 \frac{\theta}{2} \right), \quad \eta \rightarrow 0.$$

We see that I diverges logarithmically for a pure Coulomb field. As will be shown later, however, it is possible to go to the limit $\eta \rightarrow 0$ in the expression for the scattering cross section.

The differential cross section for elastic scattering of an electron can be written, according to the general formulas of Section 28, in the form

$$d\sigma_s = \frac{1}{2\pi} \sum_{\sigma_i, \sigma_f} |\bar{u}_2 S_{i \rightarrow f} u_1|^2 p_f d\phi, \quad (45.6)$$

where $S_{i \rightarrow f}$ is the sum of the matrix elements $S_{i \rightarrow f}^{(1)}, S_{i \rightarrow f}^{(2)}, S_{i \rightarrow f}^{(3)}$ with the factor $\delta(\epsilon_1 - \epsilon_2)$ omitted.

$\underline{j} = \underline{v}$ is the electron current density, (the normalizing volume is assumed equal to unity) $d\phi$ is the element of solid angle about the electron momentum after scattering and $p_f = \frac{p^2 dp}{(2\pi)^3 d\epsilon} = \frac{|p|^2}{(2\pi)^3}$

is the density of the final states per unit energy interval and unit solid angle. Equation (45.6) includes summation over electron spins in the initial and final states.

Before calculating $d\sigma_s$ let us recall that $M_1^{(2)}$ contains the photon "mass" λ which we introduced in order to avoid the infrared catastrophe. In eliminating this quantity, we must bear in mind that the probability for purely elastic electron scattering is strictly speaking zero, and that only scattering with some energy loss has physical meaning. We shall therefore add to $d\sigma_s$ the differential cross section for inelastic electron scattering with a very small energy loss $\Delta\epsilon$ (the magnitude of $\Delta\epsilon$ is assumed small compared with the electron energy $\Delta\epsilon \ll \epsilon$). The cross section for such scattering, that is the cross section for scattering of an electron with the emission of a photon whose energy is not greater than $\Delta\epsilon$, is given, as we know, by [see (32.55)]

$$d\sigma^* = \left(\frac{Ze^2}{2m\epsilon^2 \sin^2 \frac{\theta}{2}} \right)^2 (1 - v^2) \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \frac{a}{\pi} \times$$

$$\times \left\{ 2(2\Phi \coth 2\Phi - 1) \ln \frac{2\Delta\epsilon}{\lambda} + \frac{1}{v} \ln \frac{1 - v}{1 + v} - \frac{1 - v^2 \coth 2\Phi G(v, \theta)}{v \sin \frac{\theta}{2}} \right\} d\phi. \quad (45.7)$$

Let us now calculate $d\sigma_s$. On the basis of (45.3) - (45.5), the quantity $S_{i \rightarrow f}$ which enters into (45.6) can be written

$$S_{i \rightarrow f} = A\beta + B\beta\eta + C, \quad (45.8)$$

where A , B , and C do not contain Dirac matrices and are given by¹⁾

$$A = A_0 + A_1 = 2\pi l \frac{Ze^2}{q^2} - \frac{l}{2\pi} \frac{Ze^2}{q^2} \left\{ \left(1 - \frac{1}{3} \coth^2 \Phi \right) \left(1 - \frac{1}{3} \coth^2 \Phi \right) - \right.$$

$$\left. - \frac{1}{v} \right\} + (1 - 2\Phi \coth 2\Phi) \left(1 + \ln \frac{\lambda}{m} \right) + \frac{\Phi}{2} \coth \Phi + \frac{1}{2} (1 - v^2) \times$$

$$\times \frac{\coth 2\Phi}{\sin \frac{\theta}{2} \cos \frac{\theta}{2}} \int_0^1 \frac{\zeta \ln(1 - v^2 \zeta^2) d\zeta}{\zeta^2 - \cos^2 \frac{\theta}{2}} - \Phi \coth 2\Phi \ln \frac{1}{1 - v^2} \Big\} -$$

$$- 4Z^2 a^2 l m (I + J), \quad A_0 = 2\pi l \frac{Ze^2}{q^2},$$

$$B = - \frac{Ze^2}{4\pi q^2 m} \frac{\Phi}{\sinh 2\Phi},$$

$$C = - 4Z^2 a^2 l m (I - J).$$

Inserting (45.8) into (45.6) we obtain

$$d\sigma_s = \frac{1}{2\pi v} \frac{p_f}{8\pi} \text{Sp} \{ S_{i \rightarrow f} (\hat{p}_1 - m) \bar{S}_{i \rightarrow f} (\hat{p}_2 - m) \} d\phi, \quad (45.10)$$

¹⁾ In the equation for A , we have replaced $\int_0^\Phi u \tanh u du$ by the expression

$$\int_0^\Phi u \tanh u du = \frac{1}{2} \frac{1 - v^2}{\sin \frac{\theta}{2}} \frac{\coth 2\Phi}{\cos \frac{\theta}{2}} \int_0^1 \frac{\zeta \ln(1 - v^2 \zeta^2) d\zeta}{\zeta^2 - \cos^2 \frac{\theta}{2}} + \frac{\Phi}{2} \ln \frac{1}{1 - v^2}.$$

To prove this equation, let us consider the integral

$$Y = \int_{-1}^1 \frac{1 - v^2}{\sqrt{1 - p_z^2}} dx, \quad p_z = \frac{1}{2} (1 + z) p_1 + \frac{1}{2} (1 - z) p_2,$$

which enters Equation (42.4) for $M_1^{(2)}$ [see (42.16)], and let us calculate it in two ways: by performing first the substitution of variables

$$z = \frac{\tanh u}{\tanh \Phi}, \quad p_z^2 = m^2 (\cosh^2 \Phi - z^2 \sinh^2 \Phi),$$

and then

$$z = \frac{1}{\sin \frac{\theta}{2}} \sqrt{1 - \cos^2 \frac{\theta}{2}}, \quad \zeta = \frac{p_z}{p_1}.$$

Comparing the two expressions so obtained for Y , we obtain the formula given above.

where

$$\bar{S}_{i \rightarrow f} = \beta S_{i \rightarrow f}^* \beta = A^* \beta + B^* \beta q + C^*.$$

Using (28.13') it is easily shown that

$$\begin{aligned} \text{Sp} \{ S_{i \rightarrow f} (\hat{p}_1 - m) \bar{S}_{i \rightarrow f} (\hat{p}_2 - m) \} &= \\ &= 4 \{ |A|^2 (m^2 - p_1^2 p_2^2) + \text{Im} (A^* B - A B^*) + 2 m^2 (A^* C + A C^*) \} \approx \\ &\approx 4 \{ |A_0|^2 + 2 \text{Re} (A_0^* A_1) \} (m^2 - p_1^2 p_2^2) + \text{Im} (A_0^* B - A_0 B^*) + 4 m^2 \text{Re} (A_0^* C), \end{aligned} \quad (45.10')$$

where we have neglected $\bar{A}_1^2, \bar{B}_1^2, \bar{C}_1^2, \bar{B}_1 \bar{C}_1, \bar{A}_1 \bar{C}_1, \bar{A}_1 \bar{B}_1$ (the prime on the $\bar{p}$ means that the sign of the space part of the vector $\bar{p}$ should be changed). Since $\bar{A}_0 = -\frac{2\pi i Z e^2}{q^2}$ is pure imaginary, $\text{Re} (\bar{A}_0^* A_1)$ and $\text{Re} (\bar{A}_0^* C)$ are proportional to the imaginary parts of A_1 and C . From this and from (45.9) it follows that only the real parts of $\bar{A}_1 + \bar{A}_2$ and $\bar{B}_1 + \bar{B}_2$ will contribute to the scattering cross section, and according to (45.5') these do not contain the screening constant η . Therefore the final equation for the scattering cross section can be used in the case of a pure Coulomb field, as was asserted above.

Inserting Equations (45.9) and (45.10') and making use of the relation

$$m^2 - p_1^2 p_2^2 = 2e^2 \left(1 - v^2 \sin^2 \frac{\theta}{2} \right),$$

(45.10) gives the following expression for the cross section for pure elastic electron scattering:

$$\begin{aligned} d\sigma_s &= \left(\frac{Z\alpha}{2m v^2 \sin^2 \frac{\theta}{2}} \right)^2 (1 - v^2) \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) \left(1 - \frac{\alpha}{\pi} \left[2(1 - \Phi \coth \Phi) \times \right. \right. \\ &\times \left(1 - \frac{1}{3} \coth^2 \Phi \right) - \frac{2}{9} + \Phi \coth \Phi + 2(1 - 2\Phi \coth 2\Phi) \left(1 + \ln \frac{\lambda}{m} \right) + \\ &+ (1 - v^2) \frac{\coth 2\Phi}{\sin \frac{\theta}{2}} \int_0^1 \frac{\xi \ln(1 - v^2 \xi^2) d\xi}{\cos^2 \frac{\theta}{2} (1 - v^2 \xi^2) \sqrt{1 - \cos^2 \frac{\theta}{2}}} + 2\Phi \coth 2\Phi \ln \frac{1}{1 - v^2} + \\ &+ \frac{v^2 \sin^2 \frac{\theta}{2}}{1 - v^2 \sin^2 \frac{\theta}{2}} \frac{5\Phi}{\sinh 2\Phi} \left. \right] + \pi Z \alpha v \sin \frac{\theta}{2} \left(1 - \sin \frac{\theta}{2} \right) \times \\ &\times \left(1 - v^2 \sin^2 \frac{\theta}{2} \right)^{-1} \Bigg\} d\Omega, \end{aligned} \quad (45.11)$$

where $\alpha = \frac{e^2}{4\pi} = \frac{1}{137}$. We note that the term in braces which is proportional to $\frac{\alpha}{\pi}$ arises from $M_1^{(2)}$, which means that it refers to scattering in the second Born approximation; the remaining terms refer to radiative corrections (the vertex diagram). We see that the photon "mass" enters $d\sigma_s$ in the expressions $\frac{2\alpha}{\pi} (2\Phi \coth 2\Phi - 1) \ln \frac{\lambda}{m}$; in the expression for the inelastic scattering cross section $d\sigma_i$, λ enters in the term $\frac{2\alpha}{\pi} (2\Phi \coth 2\Phi - 1) \ln \frac{2\Delta\epsilon}{\lambda}$. Therefore in the total scattering cross section $d\sigma_{\Delta\epsilon} = d\sigma_s + d\sigma_i$, which

is all that has physical meaning, the photon "mass" vanishes.

We shall now write the total scattering cross section $d\sigma_{\Delta\epsilon}$ with energy losses no greater than $\Delta\epsilon$ ($\Delta\epsilon \ll \epsilon$) in the form

$$d\sigma_{\Delta\epsilon} = \left(\frac{Z\alpha}{2m v^2 \sin^2 \frac{\theta}{2}} \right)^2 (1 - v^2) \left(1 - v^2 \sin^2 \frac{\theta}{2} \right) (1 + \delta_B - \delta_R) d\Omega, \quad (45.12)$$

where δ_B and δ_R give the Born approximation and radiative corrections, respectively, and are equal to⁹⁾

$$\begin{aligned} \delta_B &= \pi Z \alpha v \sin \frac{\theta}{2} \left(1 - \sin^2 \frac{\theta}{2} \right) \left(1 - v^2 \sin^2 \frac{\theta}{2} \right)^{-1}, \\ \delta_R &= \frac{\alpha}{\pi} \left\{ 2 \left(1 - 2\Phi \coth 2\Phi \right) \left(1 + \ln \frac{2\Delta\epsilon}{m} \right) - \Phi \coth \Phi - \right. \\ &- 2 \left(1 - \Phi \coth \Phi \right) \left(1 - \frac{1}{3} \coth^2 \Phi \right) - \frac{2}{9} + \frac{1}{v} \ln \frac{1-v}{1+v} + \\ &+ \frac{2\Phi \coth 2\Phi}{1 - v^2 \sin^2 \frac{\theta}{2}} \left[\frac{v^2 \sin^2 \frac{\theta}{2}}{1 - v^2 \sin^2 \frac{\theta}{2}} - \frac{2\Phi}{\sinh 2\Phi} \right] + \\ &+ \left. \frac{(1 - v^2) \coth 2\Phi}{v \sin \frac{\theta}{2}} \int_0^1 \frac{\xi \ln(1 - v^2 \xi^2)}{1 - v^2 \xi^2} d\xi - \frac{\ln(1 - v^2)}{1 + v^2} \right\} \frac{d\Omega}{(1 - \cos^2 \frac{\theta}{2})^{1/2}}. \end{aligned} \quad (45.12')$$

We emphasize that these formulas can be used only for sufficiently small values of δ_B and δ_R . The second Born approximation gives correct results for values of Z less than about 15, and therefore Equation (45.12) is valid for sufficiently light nuclei⁹⁾. In the Table below we present values of δ_B (in per cent) for various values of the electron energy ϵ and energy loss $\Delta\epsilon$.

ϵ (MeV)	45	2.5	1.35	45	4.0	135	45	9.5	135
$\Delta\epsilon$ (deg.)	90	90	90	90	90	90	90	90	90
$\Delta\epsilon = 10$ keV	4.8	7.4	8.7	6.9	9.9	11.3	12.4	15.9	17.5
$\Delta\epsilon = 25$ "	3.9	6.0	7.1	5.7	8.1	9.3	10.5	13.5	14.8
$\Delta\epsilon = 50$ "	3.2	5.0	5.9	4.7	6.8	7.9	9.0	11.7	12.8
$\Delta\epsilon = 100$ "	2.5	3.9	4.7	3.8	5.5	6.4	7.6	9.9	10.8

We note that Equation (45.12') for δ_B becomes invalid as $\Delta\epsilon \rightarrow 0$. Since strictly speaking the scattering cross section should tend to zero as $\Delta\epsilon \rightarrow 0$, we will obtain the correct behavior for the scattering cross section as $\Delta\epsilon \rightarrow 0$, if we replace $1 - \delta_B + \delta_R$ by $e^{-\delta_B + \delta_R}$ (the other terms in the expansion of $e^{-\delta_B + \delta_R}$ in a power series of $-\delta_R + \delta_B$ will then describe higher order effects).

⁹⁾ J. Schwinger, Phys. Rev., 75, 1912 (1949); L. Elton and H. Robertson, Proc. Phys. Soc. A 65, 145 (1952).

⁹⁾ For heavier nuclei it is necessary to calculate higher Born approximations [see Urban and Mitter, Z. Naturforsch., 7a, 818 (1952)].

2. Radiative Corrections to Photon-Electron Scattering

Let us now consider radiative corrections to scattering of a free electron by a photon. Diagrams illustrating these corrections are shown in Fig. 55.

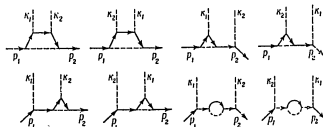


Fig. 55

The matrix elements corresponding to these diagrams can be written on the basis of the general rules of Section 24. We shall not calculate these matrix elements, but shall merely present the final results for the scattering cross section¹⁾.

In Section 29 we saw that the cross section for scattering of a photon by an electron at rest can be written in the form [see (29.20) and (29.21)]

$$d\sigma_0 = \frac{1}{2} r_0^2 \frac{x_2^2}{x_1^2} U d\Omega,$$

where r_0 is the classical electron radius, $d\Omega$ is the element of solid angle about the momentum $\mathbf{k}_2$ of the scattered photon, $\mathbf{k}_1 = \mathbf{p}_1 - \mathbf{p}_2$, $\mathbf{k}_2 = \mathbf{p}_2 - \mathbf{p}_3$, $\mathbf{k}_3 = \mathbf{p}_3 - \mathbf{p}_4$, $\mathbf{k}_4 = \mathbf{p}_4 - \mathbf{p}_5$ and

$$U = 4(x_1^{-1} + x_2^{-1})^2 - 4(x_1^{-1} + x_2^{-1}) - \left(\frac{x_1}{x_2} + \frac{x_2}{x_1}\right).$$

When we take into account radiative corrections proportional to $\frac{e^2}{\hbar^2}$, the scattering cross section becomes

$$d\sigma = \frac{1}{2} r_0^2 \frac{x_2^2}{x_1^2} \left(U - \frac{a}{\pi} U_1 \right) d\Omega, \quad (45.13)$$

where

$$U_1 = \text{Re} \{ P(x_1, x_2) + P(x_2, x_1) \}$$

¹⁾ L. Brown and R. Feynman, Phys. Rev. 85, 231 (1952).

and

$$\begin{aligned} P(x_1, x_2) = & P_A(x_1, x_2) + P_B(x_1, x_2) - (1 - 2y \cosh 2y) U \ln \lambda - \\ & - 2y \cosh 2y [2h(y) - h(2y)] U - [1 - 4y \sinh 2y (x_1 x_2)^{-1} (2 - \cosh 2y) + \\ & + 2y \cosh 2y] h(y) + \ln x_1 \left\{ 4y \cosh 2y \left[\frac{4}{x_1 x_2} \cosh^2 y + \frac{x_1 - 6}{2x_2} \cosh 2y + \right. \right. \\ & + \frac{4}{x_1^2} - \frac{1}{x_1} - \frac{x_2}{2x_1} - \frac{x_1}{x_2} - 1 \left. \right\} + \frac{3x_2}{2x_1^2} + \frac{3x_2}{2x_1} + \frac{3}{x_2} + 1 - \frac{7}{x_1 x_2} + \\ & + \frac{8}{x_1} - \frac{8}{x_1^2} + \frac{2x_1 - x_2^2 - x_1^2 x_2}{2x_1^2 (x_1 - 1)} - \frac{1}{2x_1} \frac{2x_1^2 + x_2}{(x_1 - 1)^2} + y^2 \frac{1}{\sinh^2 y} \left| \frac{2}{x_1} - \frac{7}{4} x_1 - \right. \\ & - \frac{3}{4} \frac{x_1^2}{x_1} \left. \right] - 4y \sinh y \left(\frac{1}{2} - \frac{1}{x_1} \right) + 4 \left(\frac{1}{x_1} + \frac{1}{x_2} \right)^2 - \frac{12}{x_1} - \frac{3}{2} \frac{x_1}{x_2} - 2 \frac{x_1}{x_2^2} + \\ & + \frac{1}{x_1 - 1} \left(\frac{x_1}{x_2} + \frac{1}{2} \right) + A_0(x_1) \left[\frac{x_2^2}{x_1} + \frac{x_2}{x_1^2} + \frac{x_2}{x_1} + x_1 + \frac{1}{2} x_2 + \frac{2}{x_1} - \frac{3}{x_2} - 1 \right], \end{aligned} \quad (45.14)$$

where

$$\begin{aligned} P_A(x_1, x_2) = & (1 - 2y \cosh 2y) U \ln \lambda, \quad 4 \sinh^2 y = -(x_1 + x_2), \\ h(y) = & y^{-1} \int_0^y u \cosh u \, du, \quad A_0(x_1) = -2x_1^{-1} \int_{1-x_1}^1 \ln(1-u) \frac{du}{u}. \end{aligned}$$

The quantity λ which enters the expression for P is the photon "mass". We see that the radiative corrections to photon-electron scattering diverge logarithmically as $\lambda \rightarrow 0$, as do the radiative corrections to electron scattering in an external field. In order to remove this infrared divergence, we must consider also double scattering (the double Compton effect) which involves the scattering of the electron by a photon $\mathbf{k}_1$ with the emission of two photons $\mathbf{k}_2$ and $\mathbf{k}_3$, such that the energy of the second photon is no greater than some small quantity ω_{max} much less than the energy of the electron and of photons $\mathbf{k}_1$ and $\mathbf{k}_2$. We note that this description of the problem corresponds to the usual experimental conditions, since it is impossible to guarantee that when an electron is scattered by a photon $\mathbf{k}_1$ there is one and only one photon $\mathbf{k}_2$ radiated, rather than a photon $\mathbf{k}_2$ and one or more long-wavelength photons $\mathbf{k}_3$.

The double Compton scattering cross section is given by

$$d\sigma_D = -\frac{a}{\pi} \frac{e^2}{\hbar^2} \frac{x_2^2}{x_1^2} U_D d\Omega, \quad (45.15)$$

where

$$U_D = \left\{ 2(1 - 2y \cosh 2y) \left[\ln \frac{2\omega_{\text{max}}}{\hbar} - \frac{1}{2} \right] - 4y \cosh 2y [h(2y) - 1] \right\} U. \quad (45.15')$$

The sum of the cross sections for scattering of an electron by a photon, taking into account radiative corrections and double Compton scattering is

$$d\sigma_{\Sigma} = \frac{1}{2} r_0^2 \frac{x_2^2}{x_1^2} \left[U - \frac{\alpha}{\pi} (U_1 + U_D) \right] d\phi, \quad (45.16)$$

where

$$U_1 + U_D = \text{Re} [P_0(x_1, x_2) + P_0(x_2, x_1)] + \\ + \left\{ 2(1 - 2y \coth 2y) \left(\ln 2m_{\text{max}} - \frac{1}{2} \right) + 4y \coth 2y [h(2y) - 1] \right\} U.$$

We see that the photon "mass" does not enter into this expression.

The general formulas presented above are greatly simplified in the limiting cases of low and high photon energies. For low energies, when $\omega_1 = \omega_2 = \omega \ll m$,

$$\left. \begin{aligned} d\sigma_{\Sigma} &\approx \frac{1}{2} r_0^2 d\phi (1 - 2\omega + 2\omega^2 + \dots) \left[U - \frac{\alpha}{\pi} (U_1 + U_D) \right], \\ U &\approx 1 + \cos^2 \theta, \\ U_1 &\approx -\frac{4}{3} \omega^2 (1 - \cos \theta) U \ln \lambda + (1 + \cos \theta + \cos^2 \theta - \\ &\quad - \frac{1}{3} \cos^2 \theta) 4\omega^2 \ln \omega, \\ U_D &\approx -\frac{4}{3} \omega^2 (1 - \cos \theta) U \ln \frac{2m_{\text{max}}}{\lambda} \end{aligned} \right\} \quad (45.17)$$

(θ is the photon scattering angle).

We shall consider three separate high-energy cases:

- 1) $\left| \frac{x_1 + x_2}{x_3} \right| \ll 1$,
- 2) $\left| \frac{x_1 + x_2}{x_3} \right| \sim 1$,
- 3) $\left| \frac{x_1 + x_2}{x_3} \right| \gg 1$.

In the first case

$$U = 2, \\ U_1 = 4(1 - 2y \coth 2y) \ln \lambda - 8y \coth 2y [2h(y) - h(2y)] + \\ + 4yh(y) \coth y + \ln |x_1| (4y \coth y - 1) - 2y^2 - 4y \coth y + 3 - (\ln |x_1|)^2 - \frac{\pi^2}{6}.$$

In the second case

$$U = - \left[\frac{x_1 + x_2}{x_3} \right], \\ U_1 = U \left\{ (1 - 2y) \left(\frac{3}{2} + 2 \ln \lambda \right) + 2y^2 - \frac{\pi^2}{6} \right\} + \\ + \left(1 + \frac{x_1 + x_2}{2x_3} \right) \left\{ \left[\ln \left(1 + \frac{x_2}{x_1} \right) \right]^2 - \ln \left(1 + \frac{x_2}{x_1} \right) + \right. \\ \left. + 2 \ln \frac{x_2}{|x_1|} \right\} + \left(1 + \frac{x_2}{x_1} + \frac{x_1}{2x_2} \right) \left\{ \left[\ln \left(1 + \frac{x_1}{x_2} \right) \right]^2 - \ln \left(1 + \frac{x_1}{x_2} \right) - \ln \left| \frac{x_2}{x_1} \right| + \pi^2 \right\}.$$

In the third case

$$U = - \frac{x_1}{x_2}, \\ U_1 = U \left\{ 2(1 - 2y) \ln \lambda + \ln x_3 \left[2y - \frac{3}{2} \frac{x_1 + 1}{x_2} + \frac{1}{2x_2(x_2 - 1)} \right] - \right. \\ \left. - \frac{\pi^2}{3} - Q_0(x_2) \left(\frac{1}{x_2} + \frac{y}{2} \right) + \frac{3}{2} + \frac{2}{x_2} \right\}.$$

We note that in the second and third cases $\frac{U_1}{U}$ is given by the simple expression

$$U_D = U \left\{ 2(1 - 2y) \left[\ln \frac{2m_{\text{max}}}{\lambda} - \frac{1}{2} \right] + 4y \left[y + \frac{\pi^2}{24y} - 1 \right] \right\}$$

(In the first case we must use the general equation (45.15)).

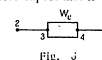
To give some idea of the order of magnitude of the radiative corrections to the Compton effect, we note that $-\frac{\alpha}{\pi} \frac{U_1}{U}$ attains its maximum value at $\theta = 0$, in which case $-\frac{\alpha}{\pi} \frac{U_1}{U} \approx 0.04$ when $\omega = 50$ Mev, and $-\frac{\alpha}{\pi} \frac{U_1}{U} \approx 0.09$ when $\omega = 1000$ Mev.

3. Natural Line Width.

In Section 34 we presented the theory of photon scattering by a bound electron. We saw that in the resonance case (when the photon energy ω is equal to the energy difference $\epsilon_n - \epsilon_j$) between electron levels the first order theory, based on the second-order S matrix, is inapplicable. The usual method for introducing radiative correction is also inapplicable in this case. Indeed, the power series expansion for the S matrix (in powers of α) is based on the assumption that each successive term is small compared with the preceding one.

This is true in the present case, when the first term of the series is infinite. It is not difficult to see that in the present case the series is in powers of the factor $-\frac{\alpha}{\epsilon_j - \epsilon_n - \omega}$, which becomes infinite for resonance. We must therefore take account of radiative corrections in an earlier stage of the calculations. (This will correspond to the fact that in the simplified derivation given in Section 34 the original wave functions were given a complex frequency.)

We note that the resonance denominator in the matrix element arises from the structure of the function ψ^E as given by (34.1), which contains $\epsilon_n + \omega$ in the denominator. We may therefore expect that if we replace ψ^E by its more accurate expression, we will obtain finite results.



Let S^E be the "exact" value of the function S^E . If S^E is represented by a solid line connecting two vertex diagrams, then the additional term $\delta S^E = S^E - S^E$ can be represented by a set of diagrams such as that of Fig. 56 with various W_E (where W_E is an electron self-energy part), i. e.,

$$\frac{1}{2} S^E(2,1) = \frac{1}{2} S^P(2,1) + \int \frac{1}{2} S^P(2,3) \Sigma(3,4) \frac{1}{2} S^P(4,1) d^4x_3 d^4x_4, \quad (45.18)$$

where $\Sigma = \Sigma(W_E)$ is a sum over all electron self-energy parts.

Equation (45.18) is a generalization of Equations (25.1)-(25.3), where δS^E was calculated in the momentum representation for the first approximation corresponding to diagram 1 of Fig. 17 (and without accounting for the external field).

Equation (45.18) can be transformed to an integral equation for S^E . To do this, we represent W_E in the form of a chain shown in Fig. 57, where each part W_E can not be broken up into parts containing only one electron line (we have called such parts proper in Section 27; the simplest W_E structures are shown in Fig. 18, diagrams 1, 3, 4, 5). The function S^E corresponds to an infinite set of chains with arbitrary numbers of W_E parts.

Therefore $\frac{1}{2} S^E$ can be written in the form of an infinite series

$$\frac{1}{2} S^E(2,1) = \frac{1}{2} S^P(2,1) + \int d^4x_3 d^4x_4 \frac{1}{2} S^P(2,3) \Sigma^*(3,4) \left\{ \frac{1}{2} S^P(4,1) + \right. \quad (45.19)$$

$$\left. + \int d^4x_5 d^4x_6 \frac{1}{2} S^P(4,5) \Sigma^*(5,6) \left[\frac{1}{2} S^P(6,1) + \dots \right] \right\},$$

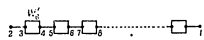


Fig. 57

where Σ^* is the sum of all Σ operators corresponding to all possible proper electron self-energy parts W_E . The series (45.19) is equivalent to the integral equation

$$\frac{1}{2} S^E(2,1) = \frac{1}{2} S^P(2,1) + \int d^4x_3 d^4x_4 \frac{1}{2} S^P(2,3) \Sigma^*(3,4) \frac{1}{2} S^E(4,1) \quad (45.20)$$

which may be written symbolically

$$\frac{1}{2} S^E = \frac{1}{2} S^P + \frac{1}{2} S^P \Sigma^* \frac{1}{2} S^E.$$

Integral equation (45.20) remains valid even when (45.19) has no meaning, as in the case of resonance.

We shall now show that S^E as found from Equation (45.20) gives a finite value for the resonance scattering matrix element which agrees with the results of Section 34. We shall attempt to find a solution to (45.20) in the following general form¹⁾:

$$\frac{1}{2} S^E = \frac{1}{2\pi i} \sum_n \int f_{nm}(w) e^{i w(t_1 - t_2)} \bar{\psi}_n(r_2) \bar{\psi}_n(r_1) d r_2, \quad (45.21)$$

¹⁾ F. Low, Phys. Rev. 88, 53 (1952).

where $\psi_n(t)$ is an electron wave function in an external field, as in Equation (34.1).

Inserting (45.21) into (45.20), we obtain the following relations for the coefficients $f_{nm}(w)$:

$$\left. \begin{aligned} f_{nn}(w) &= -\frac{1}{\epsilon_n + w} - i \frac{H_{nn}(-w)}{\epsilon_n + w} f_{nn}(w) - i \sum_{m \neq n} \frac{H_{nm}(-w)}{\epsilon_n + w} f_{nm}(w), \\ f_{nm}(w)_{n \neq m} &= -\frac{i}{\epsilon_n + w} \left\{ H_{nm}(-w) f_{nm}(w) + \sum_{p \neq m} H_{np}(-w) f_{pm}(w) \right\}, \end{aligned} \right\} \quad (45.22)$$

where

$$\left. \begin{aligned} H_{nm}(w) &= \int_{-\infty}^{\infty} H_{nm}(t) e^{i w t} dt, \\ H_{nm}(t_1 - t_2) &= \int \bar{\psi}_n(r_1) \Sigma^*(1, 2) \psi_n(r_2) d r_2. \end{aligned} \right\} \quad (45.23)$$

We note that $iH_{nn}(\epsilon_n)$ is a correction to the electron energy in state n . The $iH_{nn}(\epsilon_n)$ are generalizations of expression (26.4) for the electron self-energy to the case of motion in an external field.

If we neglect terms quadratic in H_{nm} in Equation (45.22) we obtain the following simple expressions for the f_{nm} :

$$\left. \begin{aligned} f_{nn} &= \frac{1}{\epsilon_n + w + iH_{nn}}, \\ f_{nm}(n \neq m) &= \frac{-iH_{nm}}{(\epsilon_n + w + iH_{nn})(\epsilon_m + w + iH_{mm})}. \end{aligned} \right\} \quad (45.24)$$

It is seen from this that f_{nm} (with $n \neq m$) is small compared to f_{nn} . Therefore when we insert (45.24) into (45.21), we may neglect terms containing f_{nm} with $n \neq m$. Now putting the expression obtained for S^E into the matrix $S^{(2)}$ given by (29.2), we obtain formulas which differ from those of Section 24 only in that the denominators contain $\epsilon_n + iH_{nn}$ instead of ϵ_n .

Let us separate iH_{nn} into its real and imaginary parts

$$iH_{nn} = \delta \epsilon_n - i \frac{\Gamma_n}{2},$$

where $\delta \epsilon_n$ denotes the shift of the n -th level due to radiative corrections, and Γ_n denotes the width of this level. The quantity $\delta \epsilon_n$ will be calculated by a somewhat different method in Section 46. After renormalization it is a quantity of order $\alpha^2 Z^2 \epsilon_n$ (for small Z). It should be noted that when renormalizing, a positive real quantity is subtracted. Therefore the imaginary part of iH_{nn} , that is Γ_n , should be finite even before renormalization. We shall not calculate this quantity here but merely remark that Equation (34.17) for Γ_n remains valid. The method here described can be used also to investigate higher approximations (changes in the width and shape of a resonance line).

§ 46. Radiative Level Shift for Atomic Electrons.

1. The Interaction of An Atomic Electron with the Zero-Point Field Oscillations.

The interaction of an electron with the zero-point field oscillations leads not only to additional scattering and an additional magnetic moment of the electron, but also causes shifts in atomic levels.

This can be shown from the following simple physical ideas¹⁾. Assume that the electron is moving in some static field with a potential energy $V_0(r)$. Since forces due to the zero-point field oscillations act on the electron, the position of the electron will be given by the radius vector $r + \delta r$, where δr is a displacement due to the zero-point oscillations. Therefore the potential energy of the electron will not be $V_0(r)$ but $V_0(r + \delta r)$. Since δr is small, we may write

$$V_0(r + \delta r) = \left(1 + \delta r \nabla + \frac{1}{2} (\delta r \nabla)^2 + \dots \right) V_0(r).$$

If we average this expression over all possible values of δr , we obtain the effective potential energy of the electron, namely

$$V(r) = \left[1 + \frac{1}{6} (\delta r)^2 \nabla^2 + \dots \right] V_0(r), \quad (46.1)$$

where $(\delta r)^2$ is the mean square displacement of the electron. Neglecting $(\delta r)^4$ and higher order displacements in Equation (46.1), we find that the potential energy gains an additional term

$$\delta V(r) = \frac{1}{6} (\delta r)^2 \Delta V_0(r), \quad (46.2)$$

which is proportional to the Laplacian of the original potential energy. By treating this correction as an energy perturbation, we may say that if the electron is in a state $\psi_n(r)$ with energy E_n , the level shift due to zero-point field oscillations will be

$$\delta E_n = \frac{1}{6} (\delta r)^2 \langle \psi_n | \Delta V_0 | \psi_n \rangle, \quad (46.3)$$

In particular, if we are dealing with the electron in a hydrogen atom, then

$$V_0(r) = -\frac{e^2}{4\pi r}$$

and

$$\delta E_n = \frac{2\pi}{3} (\delta r)^2 \frac{e^2}{4\pi} |\psi_n(0)|^2 \quad (46.4)$$

¹⁾ T. Welton, Phys. Rev. 74, 1157 (1948); a Russian translation of this article can be found in Atomic Electron Level Shifts (Foreign Lit. Press, 1950) p. 178.

(we have made use of the fact that $\Delta \frac{1}{r} = -4\pi \delta(r)$).

In order to find the order of magnitude of $(\delta r)^2$, we shall start with the equation of motion of an electron acted on by the zero-point field, namely

$$\ddot{r} = \frac{e}{m} E_0,$$

where E_0 is the zero-point electric field. Expanding δr and E_0 in monochromatic waves, we obtain

$$|\delta r_\omega|^2 = \frac{1}{2} \frac{e^2}{m^2 \omega^4} E_\omega^2, \quad (46.5)$$

where δr_ω and E_ω are the Fourier components, corresponding to the frequency ω , of the quantities δr and E_0 ; E_ω is given by

$$\frac{1}{2} E_\omega^2 = \frac{\omega}{2} \quad (46.5')$$

(we are assuming that the volume in which the field is located is equal to unity). From (46.5) and (46.5') it follows that

$$|\delta r_\omega|^2 = \frac{1}{2} \frac{e^2}{m^2 \omega^4} \omega.$$

Multiplying $|\delta r_\omega|^2$ by the number of frequencies in the interval $d\omega$, which is equal to $2 \frac{4\pi \omega^2 d\omega}{(2\pi)^3}$ and integrating over ω , we obtain

$$(\delta r)^2 = \frac{2}{\pi} \frac{e^2}{4\pi m^2} \int \frac{d\omega}{\omega}. \quad (46.6)$$

This expression diverges logarithmically both for low and high frequencies. If we apply arbitrary cutoffs at a low frequency of ω_0 and a high frequency of ω_1 , and if we use the expressions so obtained in (46.4), we arrive at the following formula for the displacement of the hydrogen terms²⁾:

$$\delta E_n = \frac{4}{3} \left(\frac{e^2}{4\pi m} \right)^2 \ln \frac{\omega_1}{\omega_0} |\psi_n(0)|^2. \quad (46.7)$$

²⁾ H. Bethe, Phys. Rev. 72, 339 (1947).

Thus elementary considerations show that there should exist a level shift due to the interaction of the atomic electron with the zero-point oscillations.

Equation (46.7), however, contains the undefined quantities ω_0 and ω_1 , and we shall therefore now consider the rigorous theory of this effect.

2. Radiative Atomic Level Shift.

In order to determine the radiative shift of the atomic levels, we shall make use of the effective electron potential energy introduced in Section 44, which characterizes the electron interaction with the zero-point field oscillations. Neglecting the magnetic interaction between the atomic electrons, and assuming that the atom is not in an external magnetic field, we can start with the following expression for the effective electron potential energy:

$$\delta V(x) = \frac{1}{137} \frac{e}{4\pi} \left\{ \frac{4}{3m^2} \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) \square \varphi + \frac{1}{m} \beta \alpha E \right\}, \quad (46.8)$$

where φ is the potential of the self-consistent electric field $\underline{E}$ in the atom.

Assume that the atomic electron is in a state ψ_n with energy E_n . Since the energy $\delta V(x)$ is a small perturbation, the shift of the E_n level due to interaction between the electron and the zero-point field oscillations is given by the expectation value of $\delta V(x)$ in the state ψ_n :

$$\delta E'_n = \langle \psi_n, \delta V \psi_n \rangle = \frac{1}{137} \frac{e}{4\pi} \left\{ \frac{4}{3m^2} \left(\ln \frac{m}{\lambda} - \frac{3}{8} - \frac{1}{5} \right) \langle \Delta \varphi \rangle + \frac{1}{m} \langle \beta \alpha E \rangle \right\}, \quad (46.9)$$

where $\langle L \rangle$ denotes the expectation value of the operator L in the state ψ_n .

We note, however, that like equation (46.8) this expression does not actually take into account the interaction between the electron and long-wavelength photons, so that $\delta E'_n$ gives only that part of the level shift which is due to interaction with short-wavelength photons. We shall therefore determine the shift $\delta E''_n$ due to interaction with long-wavelength photons separately. We can solve this problem by ordinary perturbation theory, using the Dirac equation for a single electron in the atom

$$i \frac{\partial}{\partial t} \psi = [\alpha(p - eA) + e\varphi + \beta m] \psi \quad (46.10)$$

and treating $U = -e\alpha A$ as the perturbation energy (A is the vector potential of the photon field, and φ is the static potential of the self-consistent atomic field).

The matrix element of this perturbation which corresponds to transition from the state $\psi_n(t) e^{-iE_n t}$ to the state $\psi_m(t) e^{-iE_m t}$ with emission of a photon of energy ω and polarization $\underline{e}$, can be written

$$U_{nm} = -\frac{e}{\sqrt{2\omega}} \langle \psi_m(r), e\alpha e^{-ikr} \psi_n(r) \rangle e^{i(E_m - E_n)t}. \quad (46.11)$$

We shall be interested in the nonrelativistic approximation. The matrix element of the perturbation energy in this case can be obtained by replacing α in (46.11) by the electron velocity operator $\underline{v} = \frac{1}{im} \nabla$. We shall

further assume that the dimensions of the atom are small compared with the wavelength $\lambda = \frac{1}{k}$, so that $e^{-ikr} \approx 1$; finally, therefore, the matrix element (which we shall write in the form $U_{nm} = \frac{V_{nm}}{i\omega} e^{i(E_m - E_n)t}$) is given by

$$V_{nm} = -\frac{e}{\sqrt{2\omega}} \langle \psi_m, (e\mathbf{v}) \psi_n \rangle. \quad (46.12)$$

Let us now consider an atomic electron in a state ψ_n with energy E_n , and let us determine the change in this energy due to the interaction between the electron and the zero-point field oscillations. This change arises in the second approximation of perturbation theory, which is as far as we shall go, and can be described in the following way: the electron emits a photon with energy ω and undergoes transition to a state ψ_m , after which it absorbs a photon and returns to the original state ψ_n . The electron energy change due to this process, as is well known, is¹⁾

$$\Delta E_n = \sum_{m, \omega, \mathbf{e}} \frac{V_{nm} V_{mn}}{E_n - E_m - \omega}, \quad (46.13)$$

where the summation is taken over all atomic electron states and over all states $(k, \mathbf{e})$ of the emitted photon, and $\frac{V_{nm}}{i\omega}$ denotes the matrix element of the perturbation energy as defined by (46.12).

Inserting (46.12) into (46.13) and replacing summation over photon states by integration over $\frac{4\pi \omega^2 d\omega}{(2\pi)^3}$ we obtain

$$\Delta E_n = \frac{e^2}{4\pi^2} \sum_{m, \mathbf{e}} \int_0^K \frac{|(e\mathbf{v})_{nm}|^2}{E_n - E_m - \omega} \omega d\omega, \quad (46.14)$$

where the integration over the photon energy is here taken from the limit $\omega = 0$ to some large energy $\omega = K$. Noting that $\sum_{\mathbf{e}} |(e\mathbf{v})_{nm}|^2 = \frac{2}{3} |\mathbf{v}_{nm}|^2$, we rewrite (46.14) in the form

$$\Delta E_n = \frac{2}{3} \frac{e^2}{4\pi^2} \int_0^K \omega d\omega \sum_m \frac{|\mathbf{v}_{nm}|^2}{E_n - E_m - \omega}. \quad (46.15)$$

For a free electron only the diagonal elements of the velocity operator are nonzero, and Equation (46.15) becomes

¹⁾ See, for instance, the quantum mechanics text books by Blokhintsev, or Landau and Lifshits.

$$(\Delta E)_{\text{free}} = -\frac{2}{3} \frac{e^2}{4\pi^2} \int_0^K \omega d\omega \frac{\mathbf{v}^2}{\omega}. \quad (46.16)$$

This expression gives the self-energy of the free electron which, as was explained in Section 26, should not be taken into account since the electromagnetic mass of the electron is already contained in the total experimentally observed electron mass. Therefore in order to find the true energy change δE_n of the bound electron, we must subtract expression (46.16) from Equation (46.15), and replace $\mathbf{v}^2$ by $(\mathbf{v})_{nm}^2$:

$$\delta E_n' = \Delta E_n - (\Delta E)_{\text{free}} = \frac{2}{3} \frac{e^2}{4\pi^2} \int_0^K \omega d\omega \left(\sum_m \frac{|\mathbf{v}_{nm}|^2}{E_n - E_m - \omega} + \frac{(\mathbf{v})_{nn}^2}{\omega} \right). \quad (46.17)$$

Since $\sum_m |\mathbf{v}_{nm}|^2 = (\mathbf{v})_{nn}^2$, Equation (46.17) can be written

$$\delta E_n' = \frac{2}{3} \frac{e^2}{4\pi^2} \int_0^K \omega d\omega \sum_m \frac{|\mathbf{v}_{nm}|^2 (E_m - E_n)}{E_m - E_n - \omega}. \quad (46.18)$$

This is the expression which gives the radiative level shifts due to the interaction between the electron and photons whose energy is no greater than K . In other words, $\delta E_n'$ is that part of the level shift which is due to the interaction between the electron and long-wavelength photons. In order to find the total level shift δE_n , we must add $\delta E_n'$ to the shift $\delta E_n''$ due to the interaction between the electron and short-wavelength photons, which is given by (46.9).

Let us perform a transformation on (46.18). We first integrate over ω . Assuming that K is large compared with all energy differences $E_m - E_n$ in the atom, we obtain

$$\delta E_n' = \frac{2}{3} \frac{e^2}{4\pi^2} \sum_m |\mathbf{v}_{nm}|^2 \ln \frac{K}{|E_m - E_n|} (E_m - E_n). \quad (46.19)$$

We now introduce the new quantity ϵ_s , defined by the formula

$$\ln \epsilon_s = \frac{\sum_m |\mathbf{v}_{nm}|^2 (E_m - E_n) \ln |E_m - E_n|}{\sum_m |\mathbf{v}_{nm}|^2 (E_m - E_n)}, \quad (46.20)$$

and rewrite $\delta E_n'$ in the form

$$\delta E_n' = \frac{2}{3} \frac{e^2}{4\pi^2} \sum_m |\mathbf{v}_{nm}|^2 (E_m - E_n) \ln \frac{K}{\epsilon_s}. \quad (46.21)$$

The sum entering this expression is well known and given by¹⁾

$$\begin{aligned} \sum_m |\mathbf{v}_{nm}|^2 (E_m - E_n) &= -\frac{1}{\pi^2} \int \psi_n^* \nabla V(r) \nabla \psi_n dV = \\ &= \frac{1}{2\pi^2} \int \Delta V(r) |\psi_n(r)|^2 dV, \end{aligned} \quad (46.22)$$

where $V(r) = e\phi$ is the potential energy of the electron. Therefore

$$\delta E_n' = \frac{e^2}{12\pi^2 m^2} e (\psi_n, \Delta \psi_n) \ln \frac{K}{\epsilon_s}. \quad (46.23)$$

Let us now consider expression (46.9) for $\delta E_n''$. We shall assume that the minimum photon energy which is accounted for in (46.9) is K , which means that it is equal to the upper limit in the integral of (46.18). In this case, according to (32.54),

$$\ln \lambda = \ln 2K - \frac{5}{6}.$$

Inserting this expression into (46.9), we obtain the following equation for the part of the level shift which is due to the interaction of the electron with photons whose energy is greater than K :

$$\delta E_n'' = \frac{e^2}{4\pi^2 \mu} e \left\{ \frac{4}{3\pi^2} \left(\ln \frac{\pi}{2K} - \frac{3}{8} + \frac{5}{6} - \frac{1}{5} \right) (\psi_n, \Delta \psi_n) + \frac{1}{m} (\psi_n, \beta \mathbf{x} E \psi_n) \right\}. \quad (46.24)$$

By adding $\delta E_n'$ and $\delta E_n''$, we find the total level shift δE_n , which in cgs electrostatic units is equal to²⁾

$$\delta E_n = \frac{1}{137} \frac{e}{4\pi} \left\{ \frac{4\pi^2}{3m^2 c^2} \left(\ln \frac{\pi}{2\epsilon_s} - \frac{3}{8} + \frac{5}{6} - \frac{1}{5} \right) \langle \Delta \varphi \rangle - i \frac{\hbar}{mc} \langle \beta \mathbf{x} E \rangle \right\}, \quad (46.25)$$

where $\langle \frac{1}{r} \rangle = (\psi_n, \frac{1}{r} \psi_n)$ and ϵ_s is defined by (46.20).

Let us determine the level shifts for the hydrogen atom. In this case (again in cgs electrostatic units)

$$V(r) = e\phi = -\frac{Ze^2}{r}$$

and

$$(\psi_n, \Delta V \psi_n) = 4\pi e^2 Z |\psi_n(0)|^2. \quad (46.26)$$

¹⁾ See H. Bethe, Quantum Mechanics of Simple Systems (United Sci. Tech. Press, 1938).

²⁾ N. Kroll and W. Lamb, Phys. Rev. 75, 388 (1949).

Further

$$|\psi_n(0)|^2 = \begin{cases} \left(\frac{Z}{na}\right)^3 \frac{1}{\pi}, & l=0, \\ 0, & l \neq 0; \end{cases} \quad (46.27)$$

$$\frac{\hbar^2}{m^2 c^2} < \Delta V > = \begin{cases} 8 \frac{a^2}{n^2} Z^4 R_y, & l=0, \\ 0, & l \neq 0; \end{cases}$$

$$\frac{e\hbar}{mc} < i\beta a E > = \begin{cases} -4a^2 \frac{Z^4}{n^2} \frac{R_y}{(l+1)(2l+1)}, & j=l+\frac{1}{2}, \\ +4a^2 \frac{Z^4}{n^2} \frac{R_y}{l(2l+1)}, & j=l-\frac{1}{2}. \end{cases}$$

where n and l are the principal and orbital quantum numbers, $a = -\frac{\hbar^2}{m e^2}$ and $R_y = \frac{1}{2} \alpha \frac{\hbar^2}{m e^2}$. Finally, for hydrogen we obtain $\frac{m}{m_e} = 7.6876$. We can use these levels to determine the shifts of the hydrogen levels whose frequencies are

$$\begin{cases} \Delta\nu(2s_{1/2}) = +1034 \text{ Mc} \\ \Delta\nu(2p_{1/2}) = -17 \text{ " } \\ \Delta\nu(2p_{3/2}) = +8 \text{ " } \end{cases} \quad (46.28)$$

The shift of the s -level is of order of magnitude $\alpha^2 E_1$, where E_1 is the energy of the ground state.

It is well known that in the ordinary Dirac theory, the $2s_{1/2}$ and $2p_{1/2}$ states have the same energy. We see that due to the interaction of the electron with the zero-point field oscillations the $2s_{1/2}$ state actually lies above the $2p_{1/2}$ state, and that the energy difference between these levels is about 1051 Mc.

Equations (46.25)–(46.27) give the radiative level shifts for the hydrogen atom up to terms of order $\alpha^2 E_1$. It can be shown that the corrections to the hydrogen level shifts, up to terms of order $\alpha^2 E_1$, are given for $n \geq 2$ state by¹⁾

$$(\delta E_n) = \frac{Z^4 a^4}{n^3} \left(1 + \frac{11}{128} - \frac{1}{2} \ln 2 + \frac{5}{192} \right) R_y. \quad (46.29)$$

When $n=2$ this gives ~ 7 Mc. The difference between the $2s_{1/2}$ and $2p_{1/2}$ levels up to terms of order α^4 is 1057.19 Mc. This result is in good agreement with the experimental value of 1057.77 ± 0.1 Mc²⁾.

3. Radiative Level Shift in Muonium.

The role of vacuum polarization is small in the radiative shift of atomic electron levels. It is therefore of interest to consider a case in which vacuum polarization is fundamental. Such a case is that of the level shift in muonium³⁾.

¹⁾ R. Karplus, A. Klein, J. and J. Schwinger Phys. Rev. 86, 288 (1952).

²⁾ Salpeter, Phys. Rev. 89, 92 (1953).

³⁾ A. Galanin and I. Pomeranchuk, Proc. Acad. Sci. USSR 86, 251 (1952).

Let us consider a negative μ -meson captured into an orbit about a proton. The lifetime of this system is sufficient for the observation of its spectrum, since it is determined by the lifetime of a free μ -meson ($2 \cdot 10^{-6}$ sec).

The radius of the normal orbit of such an "atom" will be smaller than the Compton wavelength of the electron. Since at distances of the order of the Compton wavelength of the electron the Coulomb field of the proton is altered by the polarization of the electron-positron vacuum, we may expect that the muonium spectrum will exhibit a "fine structure" whose level splitting will be of the order of $\alpha = 1/137$ of the energy of the ground state, whereas the "ordinary" fine structure is an effect of order α^2 (because of the large orbit of the electron, the part of the ordinary level shift which is due to vacuum polarization is an effect of order α^3 and is about 3 per cent of the total radiative shift).

Let $\varphi(q)$ be the Fourier transform of the external potential. According to (43.11), this potential induces a polarization potential in the vacuum given by

$$\varphi^{(p)}(q) = -\frac{\alpha}{\pi} \left[\frac{4m^2 - 2q^2}{-3q^2} (1 - \cos \theta) - \frac{1}{9} \right] \varphi(q),$$

where m is the reciprocal of the Compton wavelength of the electron, $q^2 = \mathbf{q}^2 - \omega^2$, ω is the frequency, $\mathbf{q}$ is the wave vector of the external potential, and $\cos \theta = \frac{4m^2}{4m^2 + q^2} \sin^2 \theta$. In the case of the Coulomb field

$$\varphi(q) = \frac{eZ}{q^2}$$

and

$$\varphi^{(p)}(q) = \frac{\alpha}{\pi} \frac{eZ}{q^2} \left[\frac{4m^2 - 2q^2}{3q^2} \left[1 - \frac{\sqrt{4m^2 + q^2}}{|q|} \operatorname{arctanh} \frac{|q|}{2m} \right] + \frac{1}{9} \right]. \quad (46.30)$$

Transforming to configuration space and integrating over angles, we obtain

$$e\varphi^{(p)}(r) = \frac{Ze^3}{4\pi r} \frac{2\alpha}{n^3} \int_0^\infty \left[\frac{4m^2 - 2q^2}{3q^2} \left[1 - \frac{\sqrt{4m^2 + q^2}}{|q|} \operatorname{arctanh} \frac{|q|}{2m} \right] + \frac{1}{9} \right] \frac{\sin |q|r}{|q|} d|q|. \quad (46.31)$$

Accordingly, the shift of the n -th level is given by

$$\delta E_n = -\frac{Ze^3}{4\pi} \frac{2\alpha}{n^3} \int_0^\infty \left[\frac{4m^2 - 2q^2}{3q^2} \left[1 - \frac{\sqrt{4m^2 + q^2}}{|q|} \operatorname{arctanh} \frac{|q|}{2m} \right] + \frac{1}{9} \right] \times \int \frac{\sin |q|r}{r} |\psi_n(r)|^2 dV \frac{d|q|}{|q|}. \quad (46.31)$$

Dividing δE_n by the energy of the ground state $E_1 = \frac{e^2 Z^2}{4\pi a_\mu}$ (where a_μ is the radius of the μ -meson orbit in muonium) and using $a_\mu Z$ as the unit of length, we obtain

$$\frac{\delta E_n}{E_1} = -\frac{4\pi}{3\pi^2} \int_0^\infty f(x) J_n(2\pi x) dx, \quad (46.32)$$

where

$$f(x) = \frac{1}{x^3} \left[(1-2x^2) \left(1 - \frac{\sqrt{1+x^2}}{x} \operatorname{arctanh} x \right) + \frac{x^3}{3} \right],$$

$$J_n(k) = \int_0^\infty y \sin ky R_n^2(y) dy$$

and $R_n(y)$ is the normalized radial function. The coefficient ϵ is

$$\epsilon = \frac{m}{m_\mu a Z},$$

where m_μ is the reduced mass of the μ -meson, and m is the electron mass. For the first levels we have

$$J_{10}(k) = \frac{16k}{(4+k^2)^3}, \quad J_{20}(k) = \frac{k(1-3k^2+2k^4)}{(1+k^2)^4}, \quad J_{21}(k) = \frac{k(1-k^2)}{(1+k^2)^4}.$$

Inserting the expressions for J_n into (46.32) and performing an approximate integration, we obtain the level shifts in electron volts for various Z (these are always negative):

Z	1	6	20
1s	1.8	320	20 000
2s	0.2	47	3 250
2p	0.014	27	2 550

The splitting of the $n=2$ level for $Z=1$ is 0.19 ev, which is approximately 25 times greater than the "ordinary" fine structure.

§ 47. Nonlinear Effects in Electrodynamics

1. Scattering of Light by Light

The interaction of the electromagnetic field with the zero-point oscillations of the electron-positron field leads not only to vacuum polarization, but also to several nonlinear effects in electrodynamics. Primary

among such effects is scattering of light by light.

Photon-photon scattering can be illustrated by six diagrams, one of which is shown in Fig. 58. (The other diagrams are obtained from Fig. 58 by permuting k_1, k_2, k_3, k_4 .) The scattering process represented by these diagrams can be interpreted in the following way. A photon with momentum k_1 creates a virtual pair, the electron (or positron) absorbs the photon with momentum k_2 and emits one with momentum k_3 , and finally the virtual pair annihilates into a photon with momentum k_4 . Another interpretation is also possible¹⁾; the photons k_1 and k_2 create two virtual pairs, and the electron of the first pair and positron of the second one annihilate, giving off the photon k_3 , and the positron of the first pair and electron of the second one annihilate, giving off the photon k_4 .

It is clear that from the classical point of view photon-photon scattering is a typical nonlinear effect, since according to Maxwell's equation for the electromagnetic field in vacuum, waves with different frequencies propagate independently; in other words, according to the classical electrodynamic description of the vacuum, spontaneous frequency changes are impossible.

Let us write the matrix element $S_{fi}^{(4)}$ corresponding to the diagram of Fig. 58 for the general case in which electromagnetic potentials $A_\mu(x_1), A_\mu(x_2), A_\lambda(x_3), A_\sigma(x_4)$ act at points x_1, x_2, x_3, x_4 . According to the rules of Section 24, $S_{fi}^{(4)}$ can be written



Fig. 58.

$$S_{fi}^{(4)} = -ie^4 \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{\infty} dx_3 \int_{-\infty}^{\infty} dx_4 A_\mu(x_1) A_\nu(x_2) A_\lambda(x_3) A_\sigma(x_4) \times$$

$$\times \operatorname{Sp} \left\{ \gamma_\mu \frac{1}{2} S^F(x_2 - x_1) \gamma_\nu \frac{1}{2} S^F(x_3 - x_2) \gamma_\lambda \frac{1}{2} S^F(x_4 - x_3) \gamma_\sigma \frac{1}{2} S^F(x_1 - x_4) \right\}, \quad (47.1)$$

where $\epsilon = \frac{1}{n!} = \frac{6}{4!} = \frac{1}{4}$ (here ϵ is the number of normal products of the appropriate form in the general expansion of the S -matrix into different N -products); the minus sign in (47.1) arises from the fact that the diagram contains a closed electron loop.

Transforming to the momentum representation [see (24.7)], we obtain

$$S_{fi}^{(4)} = -\frac{1}{4} \frac{e^4}{(2\pi)^4} \int d^4x \int d^4k_1 d^4k_2 d^4k_3 d^4k_4 e^{i(k_1+k_2+k_3+k_4)x} \times$$

$$\times a_\mu(k_1) a_\nu(k_2) a_\lambda(k_3) a_\sigma(k_4) T_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4), \quad (47.2)$$

where

$$T_{\mu\nu\lambda\sigma}(k_1, k_2, k_3, k_4) =$$

$$= \int d^4p \operatorname{Sp} \left\{ \gamma_\mu \frac{i\hat{p} - m}{p^2 + m^2} \gamma_\nu \frac{i(\hat{p} - \hat{k}_2) - m}{(p - k_2)^2 + m^2} \gamma_\lambda \frac{i(\hat{p} - \hat{k}_2 - \hat{k}_3) - m}{(p - k_2 - k_3)^2 + m^2} \times \right.$$

$$\times \left. \gamma_\sigma \frac{i(\hat{p} - \hat{k}_2 - \hat{k}_3 - \hat{k}_4) - m}{(p - k_2 - k_3 - k_4)^2 + m^2} \right\},$$

¹⁾ This interpretation corresponds to the diagram obtained by interchanging k_2 and k_3 in Fig. 58.

$$a_p(k) = \frac{1}{(2\pi)^3} \int A_p(x) e^{ikx} d^4x.$$

It is easily seen that $S_{i \rightarrow f}^{(0)}$ diverges logarithmically for large $|p|$, which is in agreement with the results of Section 25. According to the general rules, $S_{i \rightarrow f}^{(0)}$ is renormalized by subtracting from it its first three terms of its power series expansion in k_i , so that

$$S_{i \rightarrow f}^{(0)} = S_{i \rightarrow f}^{(0)} - S_{i \rightarrow f}^{(0)}|_{k_i = k_f = k_i = 0}. \quad (47.3)$$

We achieve the same result by subtracting from $S_{i \rightarrow f}^{(0)}$ the quantity $S_{i \rightarrow f}^{(0)}(M)$, which is obtained from $S_{i \rightarrow f}^{(0)}(m)$ by replacing the electron mass m by some large mass M , and then allowing M to approach infinity [see (27.22)¹⁾].

$$S_{i \rightarrow f}^{(0)} = \lim_{M \rightarrow \infty} [S_{i \rightarrow f}^{(0)}(M) - S_{i \rightarrow f}^{(0)}(M)]. \quad (47.3')$$

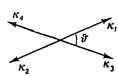


Fig. 59.

The calculation of $S_{i \rightarrow f}^{(0)}$ is performed in the same way as that of $S_{i \rightarrow f}^{(0)}$ (see Sections 42, 43). We shall not present the details of the calculation, but only the final results for the photon-photon scattering cross section for low and high frequencies. Let us consider the reference system in which the sum of the photon momenta is zero. In this system the photon frequencies are the same before and after scattering (Fig. 59). Let us denote this common frequency by ω .

It can be shown that if $\omega \ll m$, the differential scattering cross section is¹⁾

$$d\sigma = \sigma_s(\omega, \theta) d\Omega = \frac{\alpha^4}{\pi} \frac{139}{(90)^2} \frac{1}{m^2} \left(\frac{\omega}{m} \right)^6 (3 + \cos^2 \theta)^2 d\Omega, \quad (47.4)$$

where θ is the scattering angle (Fig. 59), and $d\Omega = 2\pi \sin \theta d\theta$ is the element of solid angle. The integral cross section is

$$\sigma_s(\omega) = \frac{\alpha^4}{\pi} \frac{139}{(90)^2} \frac{56}{3} \frac{1}{m^2} \left(\frac{\omega}{m} \right)^6. \quad (47.5)$$

Thus for low frequencies $\sigma_s \sim \omega^6$.

For high frequencies $\omega \gg m$, the differential scattering cross section for small angles is²⁾

$$d\sigma = \frac{\alpha^4}{\pi} \frac{1}{\omega^2} \ln^4 \theta d\Omega. \quad (47.6)$$

The integral cross section is

$$\sigma_s(\omega) = \frac{\alpha^4}{\omega^2}, \quad (47.7)$$

where α is a numerical constant.

We see that when $\omega \gg m$, the scattering cross section does not depend on the electron mass and is inversely proportional to the square of the frequency³⁾.

Since the scattering cross section is proportional to ω^6 for $\omega \ll m$ and is inversely proportional to ω^2 for $\omega \gg m$, it is a maximum roughly when $\omega \sim m$. The order of magnitude of the maximum cross section is $4 \cdot 10^{-28} \text{ cm}^2$.

The above results refer to the reference system in which the total momentum of the photon is zero. We shall now show that the scattering cross section $d\sigma$ is invariant with respect to Lorentz transformations. We note that in a time dt the photon with momentum k_1 scatters $d\sigma [1 - \cos \theta] n_2 dt$ photons with momentum k_2 , where n_2 is the density of the latter and $1 - \cos \theta$ is the relative velocity of the photons (θ is the angle between k_1 and k_2). The total number of photons with momentum k_2 scattered in a volume V during a time dt is

$$d\sigma [1 - \cos \theta] dt n_1 n_2 V,$$

where n_1 is the density of photons with momentum k_1 . This number is clearly invariant, and since $V dt$ is invariant, so is $d\sigma [1 - \cos \theta] n_1 n_2$. In other words,

$$d\sigma [1 - \cos \theta] n_1 n_2 = 2 d\sigma_0 n_1 n_2$$

and

$$d\sigma = d\sigma_0 \frac{n_1 n_2}{n_1 n_2} \frac{2}{1 - \cos \theta} \quad (47.8)$$

(the index zero denotes the reference system in which the total photon momentum is zero).

But it is known that the particle density transforms according to

$$n_0 = \frac{n - vj}{\sqrt{1 - v^2}}, \quad (47.8')$$

³⁾ It can be shown by dimensional analysis that the second of these statements follow from the first.

where $j = \frac{n \cdot k}{|k|}$ is the photon current density and $v = \frac{k_1 + k_2}{|k_1 + k_2|}$ is the relative velocity of the two reference systems in which the total photon momentum is zero and $k_1 + k_2$. Inserting (47.8) into (47.8), we obtain¹⁾

$$d\sigma = d\sigma_0, \quad (47.9)$$

so that $d\sigma$ is in fact an invariant.

In order to express σ in terms of k_1, k_2, k_3, k_4 , we can use the relation

$$(\omega_1 + \omega_2)^2 - (k_1 + k_2)^2 = 4\omega^2,$$

whence it follows that

$$2\omega^2 = \omega_1 \omega_2 [1 - \cos \theta],$$

Therefore for low frequencies

$$\sigma_s \sim \omega_1^2 \omega_2^2 [1 - \cos \theta]^3 \quad (47.10)$$

and for high frequencies

$$\sigma_s \sim \frac{1}{\omega_1 \omega_2 [1 - \cos \theta]^2} \quad (47.11)$$

2. Coherent Nuclear γ -Ray Scattering

If in Fig. 58 not all four, but only two of the dotted lines are photons and the other two represent the external electrostatic field due to a nucleus, then the diagram will represent coherent photon scattering by nuclei with no change in frequency.

It can be shown²⁾ that the integral cross section for coherent photon scattering by nuclei is given by the following formula. If $\omega \ll \frac{m}{e}$,

$$\sigma \approx (eZ)^4 r_0^2 \left(\frac{\omega}{m}\right)^4. \quad (47.12)$$

¹⁾See W. Pauli, *Helv. Phys. Acta* 6, 279 (1933).

²⁾A. Akhiezer and I. Pomeranchuk, *Physik. Z. Sowjetunion* 11, 478 (1937); H. Bethe and F. Rohrlich, *Phys. Rev.* 86, 10 (1952); F. Rohrlich and R. Glückstern, *Phys. Rev.* 86, 1 (1952).

If $\omega \gg m$,

$$\sigma \approx 0.4 (eZ)^4 r_0^2, \quad (47.13)$$

where eZ is the nuclear charge, and $r_0 = \frac{e^2}{m}$.

When $Z = 92$, Equation (47.13) gives a cross section of the order of $6 \cdot 10^{-27} \text{ cm}^2$. This is about $\frac{1}{8000}$ of the cross section for pair production by photons in the field of nucleus. Coherent photon scattering by nuclei becomes comparable to scattering by electrons at an energy of the order of 10^{18} ev . It can, however, be observed also for lower energies due to its characteristic angular distribution, which has a sharp maximum in the region of small scattering angles.

For small angles the differential scattering cross section is of the form

$$\frac{d\sigma}{d\Omega} \approx \left(\frac{Z}{2\pi}\right)^2 (Z\alpha)^4 r_0^2 \left(\frac{\omega}{m}\right)^2 F^2\left(\frac{\omega}{m}\theta\right),$$

where

$$F(x) = 0.116 - \ln x + \frac{1}{2} + \frac{x^2}{16} \dots$$

For $Z = 92$, $\omega = 300 \text{ Mev}$, and $\theta = 0.01^\circ$, this gives $3 \cdot 10^{-21} \text{ cm}^2/\text{steradian}$, whereas the Compton scattering is $7 \cdot 10^{-24} \text{ cm}^2/\text{steradian}$. When $\theta = 0^\circ$

$$\frac{d\sigma}{d\Omega} = \left(\frac{1}{32}\right)^2 \left(\frac{\omega}{m}\right)^4 (Z\alpha)^4 r_0^2.$$

3. Lagrangian Including Nonlinear Effects

The existence of nonlinear effects in electrodynamics clearly indicates that the linear Maxwell equation for the field in vacuum should be replaced by a nonlinear equation. In particular, the Lagrangian density of the classical electromagnetic field

$$L_0 = \frac{1}{2} (E^2 - H^2)$$

should be replaced by some more complicated function.

We shall assume that the field varies sufficiently slowly in space and time. This means that the change of the field $\mathbf{E}, \mathbf{H}$ over a length of order $\frac{h}{m\omega}$ and during a time interval of order $\frac{h}{m\omega}$ should be small compared with the field itself, or

$$\frac{\hbar}{mc} |\text{grad } F| \ll |F|, \quad \frac{\hbar}{mc^2} \left| \frac{\partial F}{\partial t} \right| \ll |F|. \quad (47.14)$$

Fields satisfying these conditions will not in general create real pairs, since the energy of a photon associated with such fields is much less than m^0 . We can therefore speak of a pure electromagnetic field without charged particles.

If the fields satisfy (47.14), the Lagrangian $\mathcal{L}$ will depend only on the field components, and not on their time or space derivatives. Since the Lagrangian should be invariant with respect to Lorentz transformation, the fields should enter only in the form of the two combinations $E^2 - H^2$ and $(EH)^2$, which are, as is well known, the only invariants on the field.

In the case of weak fields the Lagrangian can be expanded in powers of $E^2 - H^2$, $(EH)^2$ and we may write

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} (E^2 - H^2) + \mathcal{L}', \\ \mathcal{L}' &= \alpha (E^2 - H^2)^2 + \beta (EH)^2 + \gamma (E^2 - H^2)^3 + \delta (E^2 - H^2)(EH)^2 + \dots \end{aligned} \quad (47.15)$$

We shall show that if we know $\mathcal{S}_{1 \rightarrow f}^{(0)}$, we can find the coefficients $\alpha, \beta, \gamma, \delta, \dots$ of this expansion.

In order to establish the relation between $\mathcal{S}_{1 \rightarrow f}^{(0)}$ and $\mathcal{L}$, let us bear in mind that if the S matrix is expanded in a power series of the interaction energy between the electron-positron and electromagnetic fields, then in the first approximation we obtain

$$S_{1 \rightarrow f}^{(0)} = -i \int_{-\infty}^{\infty} V(t) dt, \quad (47.16)$$

where

$$V(t) = - \int \mathcal{L}'(x) d^3x \quad (47.16')$$

and $\mathcal{L}' = \mathcal{L} - \mathcal{L}_0$ is the difference between the Lagrangians of the coupled and free fields [see (19.13)]. Now in considering the nonlinear interactions between the electromagnetic fields, we may assume that they are described by an additional term $\mathcal{L}'$ to the Lagrangian $\mathcal{L}_0$ of the noninteracting electromagnetic fields. According to (47.16) and (47.16') this additional term should be related to $\mathcal{S}_{1 \rightarrow f}^{(0)}$ by the expression

$$S_{1 \rightarrow f}^{(0)} = i \int \mathcal{L}'(x) d^4x. \quad (47.17)$$

If we make use of this formula, then a knowledge of $\mathcal{S}_{1 \rightarrow f}^{(0)}$ is sufficient to find the coefficients $\alpha, \beta, \gamma, \delta, \dots$. We shall not go through the calculations, but shall merely present the expression for $\mathcal{L}'$ up to fourth order terms in

¹⁾ We are not considering the case of pair production by a large number of low-energy photons.

in the field¹⁾.

$$\mathcal{L}' = \frac{1}{360\pi^2} \frac{e^4}{m^4} \{ (E^2 - H^2)^3 + 7 (EH)^2 \}. \quad (47.18)$$

If we know the Lagrangian density, it is simple to find the energy density w of the field;

$$w = E \frac{\partial \mathcal{L}}{\partial E} - \mathcal{L}.$$

Inserting (47.15) and (47.18) into this expression we obtain

$$w = \frac{1}{2} (E^2 + H^2) + \frac{1}{360\pi^2} \frac{e^4}{m^4} \{ (E^2 - H^2)(3E^2 + H^2) + 7 (EH)^2 \}. \quad (47.19)$$

The first term in this expression is the classical energy density of the field, and the second term is a correction due to the nonlinear effects.

The nonlinear effects can be accounted for with the aid of a field-dependent dielectric constant and a field-dependent magnetic permeability of the vacuum. We shall determine these quantities by introducing the vectors $\mathbf{D}$ and $\mathbf{H}$ according to:

$$\mathbf{D} = \frac{\partial \mathcal{L}}{\partial \mathbf{E}}, \quad \mathbf{B} = - \frac{\partial \mathcal{L}}{\partial \mathbf{H}}. \quad (47.20)$$

From (47.15) and (47.18) we obtain

$$\left. \begin{aligned} \mathbf{D}_i &= E_i + \frac{1}{45\pi} \frac{e^4}{m^4} \{ 2(E^2 - H^2) E_i + 7 H_i (EH) \}, \\ \mathbf{B}_i &= H_i + \frac{1}{45\pi} \frac{e^4}{m^4} \{ 2(E^2 - H^2) H_i - 7 E_i (EH) \}. \end{aligned} \right\} \quad (47.21)$$

In microscopic electrodynamics the vectors $\mathbf{D}$, $\mathbf{B}$, $\mathbf{E}$, $\mathbf{H}$ are related by expressions of the form

$$\mathbf{D}_i = \epsilon_{ik} E_k, \quad \mathbf{B}_i = \mu_{ik} H_k,$$

where ϵ_{ik} and μ_{ik} are the dielectric constant and magnetic permeability tensors. Comparison with (47.21) shows that for weak and slowly varying fields the dielectric constant and magnetic permeability are given, respectively, by

²⁾ See V. Weisskopf, Kgl. Danske Videnskab. Selskab 14,1 (1936).

$$\begin{aligned} a_{ik} &= \delta_{ik} + \frac{1}{40\pi} \frac{e^2 H^2}{m^2} (2(E^2 - H^2) \delta_{ik} + 7H_i H_k), \\ \mu_{ik} &= \delta_{ik} + \frac{1}{40\pi} \frac{e^2 H^2}{m^2} (2(E^2 - H^2) \delta_{ik} - 7E_i E_k). \end{aligned} \quad (47.22)$$

A dielectric constant and a magnetic permeability which are field-dependent and different from unity lead to nonlinear effects such as refraction of light in an electric field and scattering of light by light.

Equation (47.18) gives the additional term of the Lagrangian for weak and sufficiently slowly varying fields. It can be shown that for a field which is not weak, but remains slowly varying, the additional term in the Lagrangian is given by¹⁾

$$L' = \frac{1}{8\pi^2} e^2 \int_0^\infty d\eta \left(\eta (EH) \frac{\cos \eta \sqrt{E^2 - H^2 + 2E^2 H^2}}{\cos \eta \sqrt{E^2 - H^2 + 2E^2 H^2}} + \text{c. c.} + \frac{m^4}{e^2} - \frac{\eta^2}{3} (E^2 - H^2) \right), \quad (47.23)$$

where

$$E^* = \frac{eE}{m^2}, \quad H^* = \frac{eH}{m^2}$$

and c.c. denotes the complex conjugate.

Analysis of this expression shows that for extremely strong fields the most important term in L' is of the form

$$L' \sim -\frac{e^2 E^2}{24\pi^2} \ln \frac{e|E|}{m^2}, \quad \text{or} \quad L' \sim \frac{e^2 H^2}{24\pi^2} \ln \frac{e|H|}{m^2}. \quad (47.24)$$

It follows from this that the ratio of L' to the classical Lagrangian L_0 increases only logarithmically with the field strength for very strong fields, and the coefficient of the logarithm is proportional to the fine structure constant:

$$\frac{L'}{L_0} \sim \frac{e}{3\pi} \ln \frac{e|E|}{m^2}, \quad \text{or} \quad \frac{L'}{L_0} \sim \frac{e}{3\pi} \ln \frac{e|H|}{m^2}.$$

Thus the nonlinearity in the equations of electrodynamics is only a small correction even if the field is much greater than its "critical" value $\frac{m^2}{e\hbar}$. If E is chosen as the field at the "edge of the electron", namely

$$E = \frac{m^2}{e\hbar} = 1.3 \times 10^{18} \text{ V/cm}$$

¹⁾ W. Heisenberg and H. Euler, Z. Physik 99, 714 (1936); J. Schwinger, Phys. Rev. 82, 664 (1951).

($r_0 = \frac{e^2}{4\pi m c^2}$ is the classical electron radius), then $\frac{L'}{L_0}$ becomes

$$\frac{L'}{L_0} \sim \frac{1}{3\pi} \frac{1}{137} \ln 137,$$

so that even in this case $\frac{L'}{L_0} \ll 1$.

4. Concluding Remarks

The interaction of the electromagnetic field with the electron-positron vacuum, as well as the interaction of electrons with the electromagnetic vacuum, are manifested in a remarkable way in the radiative corrections to electron scattering, the radiative atomic level shift, the anomalous magnetic moment of the electron, and in nonlinear effects in electrodynamics.

All these phenomena show that the vacuum has physical properties and cannot be considered primitive "empty" space. The concept of a physical vacuum is an extremely important feature of the present state of development of quantum electrodynamics.

The fact that there exist phenomena which are due to interactions with the vacuum is of interest not only from the purely physical but also from the general philosophical point of view. These phenomena are a new confirmation of a well-known thesis of V. I. Lenin on the inexhaustibility of the properties of the electron and the infinity of nature.

In criticizing the idealists, who concluded from the changes in concepts of the structure of matter that it can disappear, Lenin wrote "The 'essence' of things or 'substance' is also relative; it expresses only the depth of human knowledge, and because yesterday this depth extended no further than the atom, and today no further than the electron and the other, dialectical materialism insists on the temporal, relative, approximate character of all these landmarks in the understanding of nature as science progresses."

The electron is as inexhaustible as the atom, and nature is infinite; it exists forever, and this is the only categorical, the only unconditional acknowledgment of its existence outside the consciousness and perception of man, and differentiates dialectical materialism from relativistic agnosticism and from idealism." [V. I. Lenin, Materialism and Empirio-criticism (State Political Press, 1946) p.233]

The contemporary stage of development of quantum electrodynamics completely confirms these assertions of Lenin.

A great achievement of the theory is the possibility of obtaining finite values for the atomic level shift, the anomalous magnetic moment of the electron, and other quantities from divergent expressions. As we have said above, however, the procedure for removing divergences does not follow from the existing general theory and must at present be considered an auxiliary hypothesis. The fact that the radiative level shift and the anomalous magnetic moment of the electron are experimentally observed shows that this procedure is related somehow to actual physical events.

Empirical criteria confirm those results of the present state of development of quantum electrodynamics which follow from the renormalization rules. It does not follow from this, however, that these rules, as well as the general theory together with the concept of the vacuum, will not in the future undergo significant changes. On the contrary, as we have already indicated, one may assume that the difficulties of the modern theory will be eliminated only in a more complete theory which takes account of the many forms of physical interaction, including not only that set of phenomena which is related to the electromagnetic and electron-positron fields, but also the meson fields; it is entirely possible that the establishment of this more comprehensive theory will be related to further breakdowns of many fundamental physical concepts.

From the fact that concepts always change, that is, old truths assumed absolute and immutable are found to be relative, the idealists, as is well known, have concluded that objective truth would seem not to exist at all. Such a conclusion is, of course, incorrect. Knowledge is the process of the ever more complete reflection in the

human consciousness of objective natural laws. "But this," as was said by V. I. Lenin, "is not a simple, not a direct, not a complete reflection, but a process involving many abstractions, formulations, formations of concepts, laws, etc., which concepts, laws, etc., encompass conditionally and approximately the universal laws of ever moving and developing nature..." (V. I. Lenin, "Philosophical Notebooks," 1936, p. 176.)

These words of V. I. Lenin are an excellent characterization of the development of modern theoretical physics.

APPENDIX

I. THE THEORY OF WAVE FIELDS.

§ 48. The Wave Functions of the Field and Their Transformation.

1. Wave Functions of the Field and the Lorentz Group.

In the preceding chapters we have studied in detail the properties of the electromagnetic and the electron-positron fields. The theory of these most simple fields can be generalized to the case of wave fields corresponding to particles with arbitrary spins. Such a general theory of wave fields has great value in connection with the existence of various mesons which play an important role in nuclear interactions. In addition, a general theory makes possible a clearer understanding of the uniqueness of the theoretical formulations for the electromagnetic and electron-positron fields within the framework of existing physical concepts and, in particular, the uniqueness of the quantization conditions for these fields. We shall, therefore, devote this and the following sections to the consideration of some general properties of wave fields.

The field at a point $\underline{x}$ (that is, with coordinates $\underline{x}, y, z, t$) is defined by one or several functions $\psi_{\underline{i}}(\underline{x})$, which are called the components of the wave function of the field.¹⁾ The wave functions $\psi_{\underline{i}}$ satisfy certain differential equations which should be covariant with respect to Lorentz transformation; that is, their form should not change on going from one reference system to another. These equations are called the field equations.

Since concepts related to relativistic covariance play an important role in the theory of wave fields, we shall here present without proof some fundamental properties of Lorentz transformations.

Let us call the Lorentz group the set of all real transformations

$$x'_i = L_{ij} x_j \quad (i, j = 1, 2, 3, 0),$$

also written

$$x' = Lx, \quad (48.1)$$

which leave the quadratic form

¹⁾ In principle the number of components of the wave function may even be infinite.

$$s^2 = g_{ij}x_i x_j = -x^2 - y^2 - z^2 + t^2$$

$$(g_{00} = -g_{11} = -g_{22} = -g_{33} = 1, \quad g_{ij} = 0, \quad i \neq j),$$

invariant, have determinants equal to unity, and do not change the direction of time.

The Lorentz group is sometimes also called the proper Lorentz group, to distinguish it from the full Lorentz group which contains, in addition, elements corresponding to reflections in four-space. We shall speak also of the orthochronous¹⁾ Lorentz group, the set of all real transformations (48.1) which leave the four-dimensional interval invariant, have determinant of absolute value unity, and do not change the direction of time.

The components ψ_i ($i = 1, 2, 3, 4, \dots$) of the wave function are defined in every reference system, and the components $\psi'_i(x')$ in the system K' are given linearly in terms of the components $\psi_i(x)$ in the system K by an equation of the form

$$\psi'_i(x') = S_{ij} \psi_j(x), \quad x' = Lx,$$

or briefly,

$$\psi' = S\psi. \quad (48.2)$$

(The condition of linearity is related to the equivalence of inertial reference systems.) The operator S is determined by the Lorentz transformation L leading from K to K' :

$$S = S(L). \quad (48.2')$$

Since the transformation from K to K' with a subsequent transformation from K' to a third system K'' is equivalent to direct transformation from K to K'' , we may write

$$S(L'')S(L) = S(L'L), \quad (48.3)$$

¹⁾ The Russian authors call this the full or complete Lorentz group. Since, however, it does not include time reversal, this terminology is misleading. The term orthochronous has been introduced previously to describe this part of the full group [H. J. Bhabha, *Revs. Mod. Phys.* **21**, 451 (1949)].

where L' is the Lorentz transformation from K' to K'' .

Equation (48.3) shows that the set of transformations S of the components of ψ is a representation of the proper Lorentz group¹⁾ in the space of the components of the wave function. We shall call this space R , and shall call the representation D_R (the dimensionality of R is equal to the number of components of ψ and may be finite or infinite).

If we pick other components φ_i for the wave function, and if these are given by

$$\varphi_i = A_{ij} \psi_j,$$

or briefly,

$$\varphi = A\psi,$$

then the new field components will transform according to

$$\varphi'_i = A_{ij} \psi'_j = A_{ij} S_{jk} \psi_k = A_{ij} S_{jk} A_{kl}^{-1} \varphi_l,$$

or briefly,

$$\varphi' = ASA^{-1}\varphi.$$

Thus, by choosing different components of the wave function, we go from the representation $L \rightarrow S$ of the Lorentz group to the representation $L \rightarrow ASA^{-1}$. Since both of these representations are related to the same wave function, from the physical point of view they are equivalent, and shall be so-called. It is found that up to an equivalence transformation all the representations of the Lorentz group can be classified. This makes it possible both to classify all possible relativistically covariant field equations, and to establish certain general relations before securing explicit solutions of the field equations.

If there exists some choice of the components ψ_i of the wave function for which they can be divided into two classes

¹⁾ We recall that if G is a group and if each of its elements g_i ($i = 1, 2, \dots$) is put in correspondence with some operator T_{g_i} such that

$$T_{g_i} T_{g_k} = T_{g_i g_k}$$

then the mapping $g \rightarrow T_g$ is called a representation of the group G . The dimensionality of the space acted on by the operators T is called the dimensionality of the representation.

$$\psi_1, \psi_2, \dots, \psi_k \text{ and } \psi_{k+1}, \psi_{k+2}, \dots, \psi_n,$$

such that under Lorentz transformation the components of each class transform linearly among themselves, that is

$$S\psi_i = \sum_{j=1}^k S_{ij}\psi_j \quad (i=1, 2, \dots, k), \quad S\psi_i = \sum_{j=k+1}^n S_{ij}\psi_j \quad (i=k+1, \dots, n),$$

then the representation D_R is called reducible. It is clear that any reducible representation can be resolved into representations of lower dimensionality. If these latter are also reducible, then they can also be resolved into representations of even lower dimensionality. Continuing this process we finally arrive at the irreducible representations of which D_R is composed. Thus, in order to characterize a representation D_R it is sufficient to know the irreducible representations of which it is composed and the number of times any particular irreducible representation enters into it.

2. Irreducible Finite-Dimensional Representations of the Lorentz Group.

In order to find all the irreducible representations of the Lorentz group, it is sufficient to consider infinitesimal Lorentz transformations, i.e., transformations "infinitesimally close" to the identity transformation, such as

$$x'_i = x_i + \epsilon_{ij} x_j, \quad |\epsilon_{ij}| \ll 1, \quad (48.4)$$

where

$$-\epsilon_{11} = -\epsilon_{22} = -\epsilon_{33} = \epsilon_{00} = 1, \quad \epsilon_{ij} = 0, \quad i \neq j.$$

Since the interval $s^2 = g_{ij} x_i x_j$ is invariant, the matrix ϵ is antisymmetric in the first approximation, or

$$\epsilon_{ij} = -\epsilon_{ji}.$$

Therefore, an infinitesimal transformation (48.4) is defined by six parameters. These parameters may be chosen as the ϵ_{ij} , with $i > j$.

Let us now assume that in some representation D_R of the Lorentz group there is an operator $S = S(\epsilon_{ij})$ corresponding to the transformation (48.4). Expanding S in a power series in ϵ_{ij} , and restricting ourselves to the linear terms, we obtain

$$S = 1 + \sum_{i,j} \epsilon_{ij} I_{ij} = 1 + \frac{1}{2} \epsilon_{ij} I_{ij}, \quad (48.5)$$

where

$$I_{ij} = \left(\frac{\partial S}{\partial \epsilon_{ij}} \right)_{\epsilon=0}, \quad i > j; \quad I_{ij} = -I_{ji}, \quad i < j,$$

and $I = S(0)$ is the unit operator. The I_{ij} are called infinitesimal operators. These operators completely determine the representation of the group.

The infinitesimal operators are found from the following theorem; the commutator

$$[I_{ij} I_{kl}] = I_{ij} I_{kl} - I_{kl} I_{ij}$$

of two infinitesimal operators is equal to a sum of infinitesimal operators multiplied by certain coefficients c , and these coefficients are the same for all representations of the group.¹⁾

$$[I_{ij} I_{kl}] = \sum_{mn} c_{ijklmn} I_{mn}.$$

The coefficients c can be found if just one representation of the group is known. If we choose the Lorentz group itself as its own representation, we find that the infinitesimal operators of the Lorentz group satisfy the relations

$$[I_{ik} I_{jl}] = \epsilon_{il} I_{kj} + \epsilon_{kj} I_{il} - \epsilon_{il} I_{kl} - \epsilon_{kl} I_{ij} \quad (i, k, l, j = 0, 1, 2, 3). \quad (48.6)$$

These relations make it possible to find the I_{ik} . Since the representations are completely determined by the I_{ik} , in listing the representations we shall give only the formulas which define the I_{ik} .

The irreducible representations of the Lorentz group are characterized by two numbers k_0, k_1 , where k_0 is an integer or half-integer, and k_1 is an arbitrary complex number; we shall denote these representations by $\tau \sim |k_0 k_1|$.²⁾ If $2k_1$ is an integer and of the same parity (odd or even) as $2k_0$, and if $|k_1| > |k_0|$, then the

¹⁾ See V. I. Smirnov, A Course in Higher Mathematics, Vol. III, Part I.

²⁾ I. Gelfand and A. Yaglom, Izv. Akad. Nauk SSSR, Ser. Fiz.-Mat. Nauki, 1948, No. 1, p. 1.

representation will be finite dimensional; in all other cases it will be infinite dimensional. The finite dimensional representations of the Lorentz group are also characterized by two numbers j and k and shall be denoted by $\tau \sim (j, k)$. These two numbers may independently take on any of the values $j = 0, \frac{1}{2}, \dots, k = 0, \frac{1}{2}, \dots$, and are related to k_0 and K_1 by the expressions

$$j = \begin{cases} \frac{1}{2}(|k_1| + k_0 - 1), & k_1 > 0, \\ \frac{1}{2}(|k_1| - k_0 - 1), & k_1 < 0; \end{cases} \quad k = \begin{cases} \frac{1}{2}(|k_1| - k_0 - 1), & k_1 > 0, \\ \frac{1}{2}(|k_1| + k_0 - 1), & k_1 < 0. \end{cases} \quad (48.7)$$

The dimensionality of the representation $\tau \sim (j, k)$ is $(2j+1)(2k+1)$.

We shall restrict our considerations to the finite dimensional irreducible representations $\tau \sim (j, k)$, which are given in the following way. Let the ξ_m^j be a set of vectors which form a basis in $R_{(jk)}$, the space on which the representation $\tau \sim (j, k)$ acts. These vectors are determined by two numbers l and m , where m takes on the values

$$m = -l, -l+1, \dots, l,$$

and l takes on the values

$$l = |j-k|, |j-k|+1, \dots, j+k. \quad (48.8)$$

We shall now replace the ξ_{jk}^l by the operators

$$\left. \begin{aligned} H_0 &= I_{33}, \quad H_1 = I_{21} + iI_{32}, \quad H_2 = I_{21} - iI_{32}, \\ F_0 &= I_{01}, \quad F_+ = I_{03} + iI_{02}, \quad F_- = I_{03} - iI_{02}; \end{aligned} \right\} \quad (48.9)$$

and the representation $\tau \sim (j, k)$ is then defined by

$$\left. \begin{aligned} H_0 \xi_m^j &= im \xi_m^j, \\ H_+ \xi_m^j &= \sqrt{(l+m)(l-m+1)} \xi_{m-1}^j, \\ H_- \xi_m^j &= \sqrt{(l-m)(l+m+1)} \xi_{m+1}^j, \\ 2F_0 \xi_m^j &= \sqrt{(l+m)(l-m+1)} B_l \xi_{m-1}^{j-1} + \\ &\quad + m A_l \xi_m^j + \sqrt{(l+m+1)(l-m+1)} B_{l+1} \xi_{m+1}^{j+1}, \end{aligned} \right\} \quad (48.10)$$

$$\left. \begin{aligned} 2l F_+ \xi_m^j &= \sqrt{(l+m)(l-m-1)} B_l \xi_{m-1}^{j-1} + \\ &\quad + \sqrt{(l+m)(l-m-1)} A_l \xi_m^{j-1} + \\ &\quad + \sqrt{(l-m+1)(l-m+2)} B_{l+1} \xi_{m+1}^{j+1}, \\ 2l F_- \xi_m^j &= \sqrt{(l-m)(l-m-1)} B_l \xi_{m-1}^{j-1} - \\ &\quad - \sqrt{(l-m)(l-m+1)} A_l \xi_m^{j-1} + \\ &\quad + \sqrt{(l+m+1)(l+m+2)} B_{l+1} \xi_{m+1}^{j+1}, \end{aligned} \right\} \quad (48.11)$$

where

$$A_l = \frac{2(j+k+1)(j-k)}{l(l+1)}, \quad B_l = \sqrt{\frac{[l^2 - (j-k)^2][l^2 - (j+k+1)^2]}{l(l+1)}}.$$

We remark that if ψ transforms according to the representation $\tau \sim (j, k)$, then ψ^* transforms according to $\tau \sim (k, j)$.

3. Direct Product of Representations.

If the ψ_I^I and ψ_{II}^I transform according to the irreducible representations $(j_I, k_I) \sim \tau_I$ and $(j_{II}, k_{II}) \sim \tau_{II}$ respectively, then the quantities $\psi_I^I \psi_{II}^I$ also transform according to some representation of the Lorentz group; this representation is not, however, irreducible. This representation is called the direct product of the representations τ_I and τ_{II} , and is denoted by $\tau_I \times \tau_{II}$. The direct product of representations can be resolved into irreducible representations according to

$$(j_I, k_I) \times (j_{II}, k_{II}) = \sum (j, k), \quad (48.11)$$

where j and k take on the values

$$\left. \begin{aligned} j &= |j_I - j_{II}|, |j_I - j_{II}| + 1, \dots, j_I + j_{II}, \\ k &= |k_I - k_{II}|, |k_I - k_{II}| + 1, \dots, k_I + k_{II}. \end{aligned} \right\}$$

4. The Three-Dimensional Rotation Group.

An irreducible representation of the Lorentz group can in general be treated as a certain representation, in general reducible, of the rotation group in three-space. The infinitesimal rotations in three-space are given by the operators $I_{12}, I_{23},$ or I_1, I_2, I_3 . Equations (48.10) show that rotations correspond to transformations

which carry each basis vector $\epsilon_{\underline{m}}^{\underline{l}}$ into a linear combination of basis vectors with the same index $\underline{l}$. Therefore, the irreducible representations of the Lorentz group, when treated as representations of the rotation group can be resolved into $\underline{j} + \underline{k} - |\underline{j} - \underline{k}| + 1 = k_1 - k_2$ irreducible representations whose dimensionalities are $2|\underline{j} - \underline{k}| + 1$, $2(|\underline{j} - \underline{k}| + 1) + 1, \dots, 2(\underline{j} + \underline{k}) + 1$. In other words,

$$(j, k) = \sum_{l=|j-k|}^{j+k} D_l, \quad (48.12)$$

where D_l is the $(2l + 1)$ -dimensional representation of the rotation group.

The concept of the angular momentum of a quantum mechanical system is closely related to the irreducible representations of the rotation group. This relation consists of the fact that the wave functions of eigenstates of the square of the angular momentum with eigenvalues $M^2 = \hbar^2 l(l + 1)$ transform under rotation according to the irreducible representation D_l .

If two quantum mechanical systems have angular momenta $\underline{l}_1$ and $\underline{l}_2$, then as is well known, the total angular momentum of a combination of the two system can take on any of the values $\underline{l} = |\underline{l}_1 - \underline{l}_2|, |\underline{l}_1 - \underline{l}_2| + 1, \dots, \underline{l}_1 + \underline{l}_2$. This means that the wave function of the composite system transforms under a representation which is a direct product of $D_{\underline{l}_1}$ and $D_{\underline{l}_2}$:

$$D_{\underline{l}_1} \times D_{\underline{l}_2} = \sum_{l=|\underline{l}_1 - \underline{l}_2|}^{\underline{l}_1 + \underline{l}_2} D_l.$$

Comparison of this equation with (48.12) shows that the representation $\tau \sim (jk)$ of the Lorentz group, when treated as the representation of the rotation group, is equivalent to the representation according to which the wave functions of a system transform, if the system is composed of two parts with angular momenta $\underline{j}$ and $\underline{k}$.

We shall differentiate between single-valued and double-valued representations of the Lorentz group. In single-valued representations a rotation through an angle of 2π corresponds to the unit operator I , and in a double-valued representation, this rotation corresponds to the operator $-I$. It can be shown that a representation is single-valued if the numbers $2\underline{j}$ and $2\underline{k}$ have the same parity, and it is double-valued if $2\underline{j}$ and $2\underline{k}$ have opposite parity. We shall sometimes call single-valued representations tensor representations, and double valued representations spinor representations. Quantities which transform according to tensor representations are called tensors, and those which transform according to spinor representations are called spinors.

Let us present several examples. The representation $(0, 0)$ corresponds to a scalar; a vector transforms according to the representation $(\frac{1}{2}, \frac{1}{2})$; a symmetric tensor of second rank with trace zero transforms according to the representation $(1, 1)$; an antisymmetric tensor of second rank transforms according to the representations $(0, 1)$ and $(1, 0)$; a Dirac bispinor, such as those studied in Chapter II, transforms according to the representations $(0, \frac{1}{2})$ and $(\frac{1}{2}, 0)$.

¹⁾ Here $\underline{l}$ can take on integral or half-integral values.

²⁾ The sum of the diagonal elements of a tensor is invariant; if the sum is nonzero, the tensor transforms according to a reducible representation.

5. Irreducible Finite-Dimensional Representations of the Orthochronous Lorentz Group.

Let us now list all the irreducible finite dimensional representations $\underline{T}$ of the orthochronous Lorentz group. Each such representation is also a representation of the proper Lorentz group, and, therefore, it consists of the representations $\tau \sim (j, k)$. It is found that each representation $\underline{T}$ contains either one or two irreducible representations of the proper Lorentz group. In particular, if $\underline{T}$ contains $\tau \sim (j, k)$ and $j \neq k$, then $\underline{T}$ also contains the representation (k, j) . If, on the other hand, $j = k$, then $\underline{T}$ is an irreducible representation also of the proper Lorentz group. We shall denote the representations $(\underline{j}, \underline{k})$ and $(\underline{k}, \underline{j})$ which enter into $\underline{T}$ by symbols τ and $\bar{\tau}$, respectively, and we shall denote their basis vectors by $\epsilon_{\underline{m}\tau}^{\underline{l}}$ and $\epsilon_{\underline{m}\bar{\tau}}^{\underline{l}}$ (if $j = k$, then $\tau = \bar{\tau}$, and the index τ on the basis vector may be suppressed). In order to completely determine the representation $\underline{T}$, we must know the operator O corresponding to the inversion

$$x'_1 = -x_1, \quad x'_2 = -x_2, \quad x'_3 = -x_3, \quad x'_0 = x_0.$$

To do this we must distinguish between the two cases $\underline{k} \neq \underline{j}$, and $\underline{k} = \underline{j}$. In the first case O is uniquely determined by

$$O \epsilon_{\underline{m}\tau}^{\underline{l}} = (-1)^{\underline{l}} \epsilon_{\underline{m}\tau}^{\underline{l}}, \quad O \epsilon_{\underline{m}\bar{\tau}}^{\underline{l}} = (-1)^{\underline{l}} \epsilon_{\underline{m}\bar{\tau}}^{\underline{l}}, \quad (48.13)$$

where $[\underline{l}]$ is the largest integer in $\underline{l}$. In the second case O can be defined in two different ways:

$$O \epsilon_{\underline{m}\tau}^{\underline{l}} = (-1)^{\underline{l}} \epsilon_{\underline{m}\tau}^{\underline{l}}, \quad (48.14)$$

$$O \epsilon_{\underline{m}\tau}^{\underline{l}} = (-1)^{\underline{l}+1} \epsilon_{\underline{m}\tau}^{\underline{l}}, \quad (48.14')$$

and these give two nonequivalent representations.

§ 49. The Energy-Momentum Tensor and the Current Vector.

1. Energy-Momentum Tensor.

We shall assume that by treating a wave field like some generalized mechanical system, we can associate with it a Lagrangian function which is a volume integral of the Lagrangian density (or merely Lagrangian)

$L = \int \psi, \frac{\partial \psi}{\partial x_\mu}$. The Lagrangian density is a function of the field components ψ and their first derivatives with respect to space and time, and does not depend explicitly on x_μ . Then the ψ can be treated as the generalized coordinates of the system, and the $\frac{\partial \psi}{\partial x_\mu}$ as the generalized velocities.

The integral with respect to time of the Lagrangian function is called the action. In the general theory of wave fields, just as in classical mechanics and electrodynamics, we use the principle of least action according to which the action is an extremum for the actual motion of the field, i.e., for the ψ_1 which satisfy the field equations. Since the field equations are relativistically covariant, and since the action is an integral of $\mathcal{L}$ over an invariant four-dimensional volume,

$$J = \int \mathcal{L} \left(\psi_1, \frac{\partial \psi_1}{\partial x_\mu} \right) d^4x \quad (49.1)$$

(where $d^4x = dx_1 dx_2 dx_3 dx_4$), the Lagrangian density must itself be relativistically invariant.

By varying the ψ_1 and setting the variations $\delta \psi_1$ equal to zero at the boundary of the region of integration Ω , we obtain Lagrange's equations for the field

$$\frac{\partial}{\partial x_\mu} \frac{\partial \mathcal{L}}{\partial \frac{\partial \psi_1}{\partial x_\mu}} - \frac{\partial \mathcal{L}}{\partial \psi_1} = 0. \quad (49.2)$$

There are as many of these equations as there are components of the wave function.

As is well known, in classical mechanics the energy $\mathcal{E}$ of the system is related to the Lagrangian $\mathcal{L}$ by

$$\mathcal{E} = \sum_i p_i \dot{q}_i - \mathcal{L},$$

where

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

is the momentum conjugate to the coordinate q_i , and the sum is taken over all coordinates. In analogy with this, we shall define the energy density of the field by

$$w = \pi_j \dot{\psi}_j - \mathcal{L} = \pi_j \dot{\psi}_j - L, \quad (49.3)$$

where

$$\pi_j = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_j}.$$

As in mechanics, we obtain energy and momentum conservation laws. In order to show this, let us consider the second-rank tensor¹⁾

$$T_{\mu\nu} = \delta_{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial \frac{\partial \psi_1}{\partial x_\nu}} \frac{\partial \psi_1}{\partial x_\mu} = \delta_{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial \frac{\partial \psi_1}{\partial x_\mu}} \frac{\partial \psi_1}{\partial x_\nu}, \quad (49.4)$$

called the energy-momentum tensor. The energy density of the field is clearly given by

$$w = -T_{44}.$$

It can be shown that from (49.2) it follows that the four-dimensional divergence of $T_{\mu\nu}$ vanishes:

$$\frac{\partial T_{\mu\nu}}{\partial x_\nu} = 0. \quad (49.5)$$

Because the divergence of $T_{\mu\nu}$ vanishes, the four-vector P_μ , defined as the volume integral of $-iT_{\mu 4}$, is conserved:

$$P_\mu = -i \int T_{\mu 4} d^3x. \quad (49.6)$$

To see that this follows from (49.5), we write

$$\frac{dP_\mu}{dt} = \int \frac{\partial T_{\mu 4}}{\partial x_4} d^3x = - \sum_{j=1}^3 \int \frac{\partial T_{\mu j}}{\partial x_j} d^3x,$$

¹⁾ See G. Wentzel, *Introduction to the Quantum Theory of Wave Fields* (State Tech. Press, 1948); W. Pauli, *Relativistic Theory of Elementary Particles* (Harcourt Lit. Press, 1947) [the latter is a translation of an article by Pauli in *Rev. Mod. Phys.*, 13, 201 (1941)].

and since the integrand is a three-dimensional divergence, the integral can be replaced by a surface integral which may be considered zero.

The vector $\underline{p}_\mu$ is the four-dimensional energy-momentum vector of the field. Its space components form the momentum of the field $\underline{p}$, and its time component divided by c is the energy $\mathcal{E}$.

We note that the energy-momentum tensor of the field is not uniquely defined. In fact, if we add to $T_{\mu\nu}$ a quantity of the form $\frac{\partial}{\partial x_\lambda} \psi_{\mu\nu\lambda}$, where $\psi_{\mu\nu\lambda}$ can be any tensor which is antisymmetric in ν and λ , the four-dimensional divergence of $T_{\mu\nu} + \frac{\partial}{\partial x_\lambda} \psi_{\mu\nu\lambda}$ also vanishes; thus replacing $T_{\mu\nu}$ by $T_{\mu\nu} + \frac{\partial}{\partial x_\lambda} \psi_{\mu\nu\lambda}$ does not alter the total energy or momentum of the field, and therefore, this tensor may also be considered the energy-momentum tensor. In particular, the $\psi_{\mu\nu\lambda}$ may be chosen so as to make the energy-momentum tensor symmetric. A symmetric energy-momentum tensor makes possible a simple introduction of the angular momentum of the field.

We shall define the four-dimensional angular-momentum tensor of the field in the following way:

$$M_{\mu\nu} = \int (x_\nu dP_\mu - x_\mu dP_\nu) = i \int (x_\nu T_{\mu 4} - x_\mu T_{\nu 4}) d^3x, \quad (49.7)$$

where $T_{\mu\nu}$ is the symmetric energy-momentum tensor. We shall show that the components of $M_{\mu\nu}$ are conserved. To show this, it is sufficient to prove that the four-dimensional divergence of $x_\mu T_{\nu\lambda} - x_\nu T_{\mu\lambda}$ vanishes, i.e., that

$$\frac{\partial}{\partial x_\lambda} (x_\nu T_{\mu\lambda} - x_\mu T_{\nu\lambda}) = 0.$$

But this follows immediately from $\frac{\partial}{\partial x_\nu} T_{\mu\nu} = 0$, so long as $T_{\mu\nu} = T_{\nu\mu}$.

We note that the energy density $-T_{44}$ is not as a rule positive definite, and that only for certain special fields is it true that $-T_{44} \geq 0$ (compare Section 52). In all cases, however, it is necessary to satisfy the physical requirement that the total energy of the field $-\int T_{44} d^3x$ is positive.

2. Current Vector.

The field components ψ_i may be either real or complex. In the second case, we should consider the components ψ_i and their complex conjugates ψ_i^* independent.¹⁾ Then the energy-momentum tensor can be defined by

¹⁾ This means that the Lagrangian is a function both of the ψ_i and their complex conjugates ψ_i^* , and that variations of these are taken independently. The equation thus obtained for the ψ_i^* will be the complex conjugate of the equation for the ψ_i .

$$T_{\mu\nu} = L\delta_{\mu\nu} - \left(\frac{\partial L}{\partial \psi_\nu} \frac{\partial \psi_\mu}{\partial x_\nu} + \frac{\partial L}{\partial \psi_\mu^*} \frac{\partial \psi_\nu^*}{\partial x_\mu} \right). \quad (49.8)$$

We shall assume that in the case of complex fields, the Lagrangian is invariant under the transformation

$$\psi \rightarrow \psi' = \psi e^{i\alpha}, \quad \psi^* \rightarrow \psi'^* = \psi^* e^{-i\alpha}, \quad (49.9)$$

here α is some arbitrary real quantity. It then follows that

$$\frac{1}{i} \left(\frac{dL}{d\alpha} \right)_{\alpha=0} = \frac{\partial L}{\partial \psi_\nu} \psi_\nu + \frac{\partial L}{\partial \psi_\nu^*} \psi_\nu^* - \frac{\partial L}{\partial \psi_\nu} \psi_\nu^* - \frac{\partial L}{\partial \psi_\nu^*} \psi_\nu = 0. \quad (49.9')$$

(We therefore define the four-vector²⁾

$$J_\mu = ie \left(\frac{\partial L}{\partial \psi_\mu} \psi - \psi^* \frac{\partial L}{\partial \psi_\mu^*} \right). \quad (49.10)$$

where e is some constant, then according to (49.9') and Lagrange's equations (49.2),

$$\frac{\partial J_\mu}{\partial x_\mu} = 0. \quad (49.11)$$

We shall interpret J_μ as the electric-current density of the field. When the quantum conditions are applied to the field, this interpretation leads to the charge being an integral multiple of e (compare Section 17). Therefore, e is considered the charge of the individual particles associated with the field.

In the case of real fields, the current vector vanishes. Therefore, real fields, which do not admit the transformation (49.9), describe neutral particles. Charged particles are always described by complex fields.

50. Relativistically Covariant Field Equations.

1. General Form of Relativistically Covariant Field Equations.

Let us consider noninteracting fields for which the superposition principle is valid, and which, therefore, satisfy linear differential equations. In view of the homogeneity of space and time, these equations must have constant coefficients.

¹⁾ W. Pauli. See the reference on p. 505.

We may assume without loss of generality that the differential equations of a free (noninteracting) field are first-order homogeneous equations, since higher order equations can be reduced to first-order equations by considering the derivatives of the wave function as field components. We shall therefore start with equations of the form

$$\Gamma_\mu \frac{\partial \psi}{\partial x_\mu} + \kappa \psi = 0, \quad (50.1)$$

where ψ is the set of field components, Γ_μ is a matrix acting on component indices, and κ is a real constant differing from zero.¹⁾ The problem under consideration is the following: we wish to find all possible equations of the form of (50.1) which are invariant with respect to Lorentz transformations.²⁾

It should be borne in mind that a given equation of the form (50.1) can be satisfied by wave functions which describe different fields (for instance, a scalar and a pseudoscalar). Therefore, a field is characterized not only by the Γ_μ matrices in Equation (50.1), but also by the representation D_R according to which the wave functions ψ transform.

The derivatives $\frac{\partial \psi}{\partial x_\mu}$ transform according to the direct product of D_R and $\tau \sim \left(\frac{1}{2}, \frac{1}{2}\right)$, the latter being the representation according to which the four-vector x transforms. It is seen from Equation (50.1) that the linear combinations $\Gamma_\mu \frac{\partial \psi}{\partial x_\mu}$ transform according to the same representation D_R as do the wave functions ψ . In other words, the representation D_R is contained in the direct product $D_R \times \left(\frac{1}{2}, \frac{1}{2}\right)$, or

$$D_R \subset D_R \times \left(\frac{1}{2}, \frac{1}{2}\right), \quad (50.2)$$

and the Γ_μ matrices project the space of the derivatives $\frac{\partial \psi}{\partial x_\mu}$ onto the space R of the wave functions ψ .

Expression (50.2) shows that if an irreducible representation $\tau \sim (j, k)$ of the Lorentz group is contained in D_R , then according to (48.11) D_R contains also at least one of the following four irreducible representations:

$$\tau \sim \left\{ \begin{aligned} &(j - \frac{1}{2}, k - \frac{1}{2}), (j - \frac{1}{2}, k + \frac{1}{2}), \\ &(j + \frac{1}{2}, k - \frac{1}{2}), (j + \frac{1}{2}, k + \frac{1}{2}) \end{aligned} \right\} \quad (50.3)$$

¹⁾ The case $\kappa = 0$, as for the electromagnetic field, is considered in Section 15. The requirement that κ be real is no restriction on the generality of the considerations. In fact, if we choose κ complex in (50.1), and multiply this equation by $\frac{1}{\kappa}$, we obtain an equivalent equation with κ real.

²⁾ The treatment presented here is due to I. Gelfand and A. Yaglom, J. Exptl.-Theoret. Phys. 18, 763 (1948).

Inserting (48.1) and (48.2) into (50.1) leads to the following condition which must be satisfied by (50.1) under Lorentz transformations:

$$\Gamma_\mu := L_\mu S \Gamma_\nu S^{-1}. \quad (50.4)$$

This condition must be satisfied, in particular, for the infinitesimal Lorentz transformations defined by (48.4) and (48.5). With these equations and (50.4) we obtain

$$[\Gamma_\mu, \Gamma_\nu] = g_{\mu\nu} \Gamma_4 - g_{\mu 4} \Gamma_\nu, \quad (50.5)$$

where the Γ_{ij} are given by (48.10) and (48.9). It follows from (50.5) that

$$\Gamma_k = (\Gamma_0, \Gamma_{0k}), \quad k = 1, 2, 3, \quad (50.6)$$

and therefore, the problem reduces to determining all possible operators Γ_0 .

The representation D_R according to which the components of ψ transform can be resolved into irreducible representations $\tau \sim (j, k)$. We shall thus denote the basis vectors in the space R of the wave functions ψ by $\epsilon_{m\tau}^j$, in order to indicate the irreducible representation τ according to which they transform. The operator Γ_0 can clearly be given by relations of the form

$$\Gamma_0^j \epsilon_{m\tau}^j = \sum_{m', \tau'} c_{m\tau, m'\tau'}^j \epsilon_{m'\tau'}^j, \quad (50.7)$$

where c are the desired matrix elements of Γ_0 . Inserting $i = 0, j, k = 1, 2, 3$ into (50.5), it is easily shown that the operator Γ_0 commutes with H_0, H_+ and H_- , or with all the transformations of that representation of the three-dimensional rotation group which is defined by the representation D_R of the Lorentz group. It follows from this that the matrix c should be diagonal with respect to the indices i, l and m, m' , i.e.,

$$c_{m\tau, m'\tau'}^j = c_{m\tau}^j \delta_{m\tau, m'\tau'},$$

and, therefore, (50.7) reduces to

$$\Gamma_0^j \epsilon_{m\tau}^j = c_{m\tau}^j \epsilon_{m\tau}^j. \quad (50.8)$$

The coefficients $c_{m\tau}^j$ are different from zero only if each of the numbers of the pair (j, k) which determines the representation τ differ by $\pm \frac{1}{2}$ from the corresponding number of the pair (j, k) which determines τ .

that is, if $\tau' \sim (j', k')$ is one of the representations of (50.3). If this is the case, there are two possibilities; if $(j', k') = \left(j + \frac{1}{2}, k - \frac{1}{2} \right)$, then

$$\begin{aligned} c_{\tau'}^j &= c_{\tau'} \sqrt{(l+j-k+1)(l-j-k+1)}, \\ c_{\tau'}^k &= c_{\tau'} \sqrt{(l+j-k+1)(l-j-k+1)}, \end{aligned} \quad (50.9)$$

if $(j', k') = \left(j + \frac{1}{2}, k + \frac{1}{2} \right)$, then

$$\begin{aligned} c_{\tau'}^j &= c_{\tau'} \sqrt{(l+j+k)(l-j-k+1)}, \\ c_{\tau'}^k &= c_{\tau'} \sqrt{(l+j+k)(l-j-k+1)}, \end{aligned} \quad (50.9')$$

where $c_{\tau'}$, and $c_{\tau'}$ are arbitrary complex numbers.

We shall call the representations τ and τ' linked if $c_{\tau'} \neq 0$. It follows from the above that the representations $\tau \sim (j, k)$ and $\tau' \sim (j', k')$ may be linked if

$$|j' - j| = \frac{1}{2}, |k' - k| = \frac{1}{2}. \quad (50.10)$$

In order that equation (50.1) be invariant not only under the proper Lorentz group, but also under the inversion θ

$$x_0 \rightarrow x_0, \quad x_1 \rightarrow -x_1, \quad x_2 \rightarrow -x_2, \quad x_3 \rightarrow -x_3,$$

the $c_{\tau'}$ must satisfy certain conditions. These conditions are the following. If τ and τ' denote the representations (j, k) and (j', k') , these numbers must satisfy the relations

$$\begin{aligned} c_{\tau'}^j &= c_{\tau'}^j, & \tau &\neq \tau', & \tau' &\neq \tau', \\ c_{\tau'}^k &= c_{\tau'}^k, & \tau &= \tau', & \tau' &\neq \tau', \\ c_{\tau'}^j &= c_{\tau'}^j, & \tau &= \tau', & \tau' &= \tau'. \end{aligned} \quad (50.11)$$

[the positive or negative sign is chosen depending on whether θ is given by (48.14) or (48.14')].

d) Let us recall that the proper Lorentz group and this inversion together form the orthochronous Lorentz group.

Equations (50.6), (50.8), (50.9), (50.9'), and (50.11) are a complete solution to the problem of finding all possible Lorentz covariant equations.

2. Invariant Lagrangian.

It is not in general possible to relate such physical quantities as the energy, momentum, angular momentum, charge, etc., to Equation (50.1). These quantities can be defined if Equation (50.1) corresponds to a real invariant Lagrangian.

A real invariant Lagrangian leading to Equation (50.1) can always be written in the form

$$L = \frac{1}{2i} \left(\Gamma_{\mu} \frac{\partial \psi}{\partial x_{\mu}} \right) - \left(\psi, \Gamma_{\mu} \frac{\partial \psi}{\partial x_{\mu}} \right) + \kappa (\psi, \psi), \quad (50.12)$$

where (ψ_1, ψ_2) is some invariant nondegenerate bilinear form of the wave functions ψ_1 and ψ_2 . A bilinear invariant form is defined as follows. If we are given the expressions for ψ_1 and ψ_2 in terms of the basis vectors $\psi_{m\tau}^j$, namely

$$\psi_1 = \sum_{m\tau} x_{m\tau}^j \psi_{m\tau}^j, \quad \psi_2 = \sum_{m\tau} y_{m\tau}^j \psi_{m\tau}^j,$$

then the bilinear form is given by

$$(\psi_1, \psi_2) = \sum_{m\tau, m'\tau'} a_{m\tau, m'\tau'}^j x_{m\tau}^j y_{m'\tau'}^j, \quad (50.13)$$

where the quantities

$$a_{m\tau, m'\tau'}^j = (c_{m\tau}^j, c_{m'\tau'}^j)$$

must be given.

A bilinear form is called nondegenerate if there exists no wave function ψ_0 for which $(\psi_0, \varphi) = 0$ for all φ . The invariance of the form means that the relation

$$(S\psi_1, S\psi_2) = (\psi_1, \psi_2) \quad (50.14)$$

is true for all the transformations S of the function ψ , which make up the representation $D_{\frac{1}{2}}$ of the Lorentz group. In other words, (50.13) transforms under the unit representation of the Lorentz group.

A product of the form $\sum_{m\tau} x_{m\tau}^j \psi_{m\tau}^j$ transforms under the direct product of the representations $\tau \times X_{\frac{1}{2}}$ [we recall that if $\tau' \sim (j', k')$, then $\tau' \sim (k', j')$]. According to (48.11), $\tau \times X_{\frac{1}{2}}$ contains the unit representation $\tau \sim (0, 0)$ of the Lorentz group only if $k' = j$ and $j' = k$, that is if $\tau' = \tau$. Therefore, if $\tau' \neq \tau$, then

the $a_{\overline{m}\tau, \overline{m}'\tau'}^{II'} = 0$. Applying similar considerations to the three-dimensional rotation subgroup, it can be shown that the $a_{\overline{m}\tau, \overline{m}'\tau'}^{II'}$ are of the form

$$a_{\overline{m}\tau, \overline{m}'\tau'}^{II'} = a_{\tau\tau'}^{II'} \delta_{\overline{m}\overline{m}'} \delta_{\tau\tau'}, \quad (50.15)$$

where

$$a_{\tau\tau'}^{II'} = (-1)^{II'} a_{\tau\tau'}, \quad (50.15')$$

and the $a_{\tau\tau'}$ are numbers which satisfy

$$a_{\tau\tau'} = a_{\tau\tau'}^*, \quad (50.16)$$

and are otherwise arbitrary.

Let us find the kind of equations we are led to by the Lagrangian (50.12). To do this we set the variation of the action

$$I = \int L d^4x,$$

equal to zero. The variation of I is

$$\delta L = \frac{1}{i} \left\{ \left(\frac{1}{2} (\Gamma_k + \Gamma_k^*) \frac{\partial \psi}{\partial x_k} + i\psi \right) \delta \psi - \left(\delta \psi, \frac{1}{2} (\Gamma_k + \Gamma_k^*) \frac{\partial \psi}{\partial x_k} + i\psi \right) \right\} + \frac{\partial}{\partial x_k} \frac{1}{2i} \{ (\Gamma_k \delta \psi, \psi) - (\psi, \Gamma_k \delta \psi) \}, \quad (50.17)$$

where Γ_k^* is given by the relation

$$(\Gamma_k \psi, \psi) = (\psi, \Gamma_k^* \psi). \quad (50.17')$$

Integrating (50.17) over a four-dimensional volume and dropping the last term, we obtain the following Lagrange's equation:

$$\frac{1}{2} (\Gamma_k + \Gamma_k^*) \frac{\partial \psi}{\partial x_k} + i\psi = 0. \quad (50.18)$$

This is the same as (50.1) if

$$\Gamma_i = \Gamma_i^* \quad (i = 0, 1, 2, 3, 4),$$

i.e., if for arbitrary ψ_1 and ψ_2

$$(\Gamma_i \psi_1, \psi_2) = (\psi_1, \Gamma_i \psi_2). \quad (50.19)$$

Equations (50.19) show that the Γ_i are self-conjugate with respect to the bilinear form (ψ_1, ψ_2) . Thus, in order that Equation (50.1) be a Lagrange's equation, there must exist a bilinear invariant form for which the operators Γ_i are self-conjugate.

The property of the Γ_i places an additional condition on their choice, that is, on the choice of the $c_{\tau\tau'}$. It can be shown from the fact that Γ_0 is self-conjugate, that

$$c_{\tau\tau'} a_{\tau'\tau} = c_{\tau'\tau}^* a_{\tau\tau'}. \quad (50.20)$$

This is the final condition that the matrix Γ_0 must satisfy in order that Equation (50.1) be obtainable from an invariant Lagrangian.

We note that if we transform to a new basis according to

$$\xi_{\overline{m}\tau}^i = \sum_{\tau'} \alpha(\tau, \tau') \xi_{\overline{m}\tau'}^i, \quad \xi_{\overline{m}\tau}^i = \sum_{\tau'} \alpha(\tau', \tau) \xi_{\overline{m}\tau'}^i,$$

(the sum is taken over all representations τ' in $\underline{D}_R$ which are equivalent to τ), it is possible to choose the quantities α so that

$$\alpha_{\tau\tau'} = \pm \delta_{\tau\tau'}. \quad (50.21)$$

When this is done the difference between the number of positive and negative coefficients is independent of the choice of $\alpha(\tau, \tau')$, and is determined by the initial $a_{\tau\tau'}$ matrix. Thus, two bilinear forms of the type of (50.21) with a different number of positive coefficients cannot be transformed one into the other, and are thus not physically equivalent.

The Lagrangian (50.12) corresponds to the energy momentum tensor

$$T_{\mu\nu} = \frac{1}{2i} \left\{ \left(\Gamma_\mu \frac{\partial \psi}{\partial x_\nu} - \psi, \Gamma_\nu \frac{\partial \psi}{\partial x_\mu} \right) \right\} \quad (50.22)$$

and the current vector

$$j_\mu = e(\Gamma_\mu \psi, \psi). \quad (50.23)$$

§ 51. The Mass and Spin of a Particle.

1. Mass and Spin Values for a Given Equation.

As we shall see later, the way a field which transforms under a finite-dimensional representation of the Lorentz group must be quantized depends on whether the representation is single-valued or double-valued. It is therefore, of interest to establish first of all which of the properties and quantities that characterize the particles can be found in general directly from the relativistically covariant Equation (50.1). These quantities are the mass and spin, or more exactly the masses and spins, of the particles associated with the wave field described by (50.1); it is possible in general that these particles may have states with different mass and spin values.

In order to find all possible values of the mass and spin of the particles described by (50.1), it is sufficient to consider the plane-wave solutions of these equations in the rest system of the particle,

$$\psi(x) = \psi^{(0)} e^{-ip_0 t}, \quad (51.1)$$

where p_0 is the rest energy of the particle, and $\psi^{(0)}$ is a constant. When (51.1) is inserted into (50.1), we obtain

$$p_0 \Gamma_0 \psi^{(0)} - x \psi^{(0)} = 0, \quad (51.2)$$

whence

$$p_0 = \frac{x}{\lambda_1}, \quad (51.3)$$

where λ_1 is some nonzero eigenvalue of Γ_0 . This shows that the possible mass values of the particles corresponding to (50.1) are

$$m_i = \frac{\hbar}{c} \frac{x}{\lambda_i}, \quad (51.3)$$

where λ_i runs through all possible eigenvalues of the matrix Γ_0 .

2. The Spin Values of Particles Described by a Given Equation.

Let us now determine the possible values for the spin of the particles described by (50.1). We note that Γ_0 can be broken up into separate parts $\| \mathbb{C}_{TT}^I \|$ corresponding to different values of I , and that the matrix $\| \mathbb{C}_{TT}^I \|$ enters into Γ_0 exactly $2I + 1$ times. Each eigenvalue $\lambda_i^{(I)}$ of the matrix $\| \mathbb{C}_{TT}^I \|$ is a $(2I + 1)$ -fold eigenvalue of Γ_0 . The $2I + 1$ eigenfunctions of Γ_0 which belong to this eigenvalue transform among themselves

under three-dimensional rotations according to some irreducible representation D_I of the rotation group. These $2I + 1$ eigenfunctions of Γ_0 correspond to different states of the particle with mass $\hbar/\lambda_i^{(I)}$ and spin I , these states differing in the projection of the spin along some axis.

Thus, the possible spin values of the particles described by (50.1) are given by those values of I for which the matrix $\| \mathbb{C}_{TT}^I \|$ has nonzero eigenvalues, and the number of these eigenvalues is the number of different states with spin I , which in general differ in their mass.

Since I varies from $|j - k|$ to $j + k$, the spin will be integral if both j and k are integers or odd half-integers, i.e., if the representation (j, k) is single-valued. If one of the pair j, k is integral, and the other is half-integral, that is, if the representation of (j, k) is double-valued, the spin is half-integral.

This definition of the spin assumes that the mass of the particle is not zero (otherwise, there is no rest system). However, the concept of spin exists also for particles of mass zero. In this case, the spin I can be defined as the minimum value of the angular momentum of the field which is always equal to $\sqrt{I(I+1)}$.

Thus, the possible values of the mass and spin of the particles are directly related to the matrix Γ_0 ; the masses are given by the eigenvalues, and the spins by the multiplicities of the eigenvalues of this matrix.

Related to this we may also ask how to find the form of the relativistically covariant Equation (50.1) corresponding to a particle with given mass and spin. This problem reduces to finding Γ_0 , i.e., the quantities $\mathbb{C}_{TT}^I$, from the given values of m and I . We shall show in the following section how to solve to this type of problem under certain additional assumptions.

We note that a field which transforms according to an infinite-dimensional representation of the Lorentz group also describes particles which can be in states with various mass values, and that as a rule the mass spectrum approaches zero as $I \rightarrow \infty$, and only in exceptional cases is the mass independent of I for sufficiently large I .

§ 52. Examples of Wave Equations for Particles of Various Spins.

1. Wave Equation with Positive Definite Charge Density.

Let us consider several applications of the general theory of relativistically covariant wave equations. Let us first pose the following problem; we wish to find an equation of the form of (50.1) which describes particles with positive definite charge density.

Since the charge density is given by

$$\rho = e(\Gamma_0 \psi, \psi),$$

the problem reduces to finding all equations of the form of (50.1) whose solutions satisfy the condition

$$(\Gamma_0 \psi, \psi) \geq 0. \quad (52.1)$$

This condition, should, in particular, be satisfied by plane waves in the rest system of a particle, i.e., solutions of the form $\psi = \psi^{(0)} e^{-ip_0 t}$, where $\psi^{(0)}$ is an eigenvector of Γ_0 ; thus,

$$(\Gamma_0 \psi, \psi) \geq 0. \quad (52.2)$$

We shall assume that this matrix can be diagonalized. In this case the set of eigenvectors $\psi^{(s)}$ of the matrix Γ_0 is complete, and therefore, an arbitrary vector ψ can be expanded in terms of the eigenvalues of Γ_0 . Thus, Equation (52.1) will be satisfied for arbitrary ψ , and not only for solutions of (50.1). Let us use this fact and write an arbitrary vector ψ in the form

$$\psi = \sum_{lm} x_{lm}^i \frac{t^i}{r^i}.$$

The positive form $(\Gamma_0 \psi, \psi)$ can, according to (50.8), (50.13), (50.15), and (50.15'), be written

$$(\Gamma_0 \psi, \psi) = \sum_{lm} (-1)^l a_{lr} c_{lr} \eta(l) x_{lm}^i x_{lm}^i > 0, \quad (52.3)$$

where $\eta(l) = \frac{c_{lr}}{r^l} / \frac{c_{lr}}{r^l} > 0$ [see (50.9)]. A necessary condition for $(\Gamma_0 \psi, \psi)$ to be a positive form is that the coefficients of the $\left| \frac{x_{lm}^i}{r^l} \right|^2$ be positive;

$$(-1)^l \sum_{lr} a_{lr} c_{lr} > 0. \quad (52.3')$$

This necessitates that there exist a pair of representations τ and τ' for which $a_{\tau\tau'} \neq 0$ and $c_{\tau\tau'} \neq 0$. The first of these conditions, according to (50.15), means that $\tau' = \tau$, and therefore, the second condition means that τ and τ' are linked representations [see (50.11)]. In other words, if $\tau \sim (j, k)$ and $\tau' \sim (k, j)$, then $k = j + 1/2$ or $k = j - 1/2$. Let us set $\tau \sim (j, j - 1/2)$, $\tau' \sim (j - 1/2, j)$ and we can then show that $j = 1/2$. This follows from the fact that according to (48.8) l takes on the values $l = 1/2, 3/2, \dots, 2j - 1/2$, and for $j > 1/2$ the factor $(-1)^l$ in (52.3) would change sign and condition (52.3') would not be satisfied.

Thus, we have shown that the representation D_R under which ψ transforms breaks up into $\tau \sim (1/2, 0)$ and $\tau' \sim (0, 1/2)$. These representations correspond to spin $1/2$.

It can be shown that if a representation were to contain several pairs of representations $\tau \sim (1/2, 0)$, $\tau' \sim (0, 1/2)$, Equation (50.1) would break up into several independent invariant equations. Therefore, D_R contains the representations $\tau \sim (1/2, 0)$, $\tau' \sim (0, 1/2)$ only once, and is thus a four-dimensional representation.

Let us now find the matrix Γ_0 which is given by the relation

$$\Gamma_0 \frac{t^i}{r^i} = \sum_{lr} c_{lr}^i \frac{t^l}{r^l}.$$

Since neither $\tau \sim (1/2, 0)$ nor $\tau' \sim (0, 1/2)$ is linked with itself, $c_{\tau\tau} = c_{\tau'\tau'} = 0$. Further, inserting $\tau' = \tau$ into (50.20), we find that

$$c_{\tau\tau}^i a_{\tau\tau} = a_{\tau\tau}^i c_{\tau\tau}^i,$$

and since according to (50.16) $a_{\tau\tau} = a_{\tau'\tau'}$, we have

$$c_{\tau\tau}^i = c_{\tau'\tau'}^i = c_{\tau\tau}^i.$$

[The last equation follows from (50.11) with $\tau' = \tau$]. Without loss of generality we set $c_{\tau\tau} = 1$ [this is always possible by dividing Equation (50.1) by $c_{\tau\tau}$].

Thus, $c_{\tau\tau} = c_{\tau'\tau'} = 0$, $c_{\tau\tau'} = c_{\tau'\tau} = 1$. With these relations and Equation (50.9) it is easily shown that Γ_0 is of the form

$$\Gamma_0 = \begin{pmatrix} \frac{t^{1/2}}{r^{1/2}} & \frac{t^{1/2}}{r^{1/2}} & \frac{t^{1/2}}{r^{1/2}} & \frac{t^{1/2}}{r^{1/2}} \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (52.4)$$

(the symbols above the matrix show the way the basis vectors $\frac{t^i}{r^i}$ are ordered).

Let us now find the remaining matrices Γ_k (with $k = 1, 2, 3$) in Equation (50.1). These are given by Equations (50.6)

$$\Gamma_k = i \Gamma_0 \Gamma_{0k}, \quad k = 1, 2, 3,$$

where Γ_{0k} can be found with the aid of (48.10), namely

$$F_0 = \Gamma_{01} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix},$$

$$F_+ = I_{03} + iI_{02} = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad F_- = I_{03} - iI_{02} = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

It then follows that

$$I_{03} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad I_{02} = \frac{i}{2} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Inserting these matrices in (50.6), we obtain

$$\Gamma_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \Gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \Gamma_3 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \quad (52.5)$$

Introducing the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

we can write $\Gamma_0, \Gamma_1, \Gamma_2, \Gamma_3$ in the form

$$\Gamma_0 = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}, \quad \Gamma_1 = \begin{pmatrix} -i\sigma_y & 0 \\ 0 & i\sigma_y \end{pmatrix}, \\ \Gamma_2 = \begin{pmatrix} 0 & -i\sigma_y \\ i\sigma_y & 0 \end{pmatrix}, \quad \Gamma_3 = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}.$$

It is easily shown that the Γ_i satisfy the relations

$$\Gamma_i \Gamma_k + \Gamma_k \Gamma_i = 2\delta_{ik}, \quad i, k = 0, 1, 2, 3. \quad (52.6)$$

These relations show that the Γ_i -matrices are equivalent to the Dirac γ_i -matrices and therefore, in the present case Equation (50.1) is the Dirac equation which we have studied in Section 7.

We have thus shown that of all the finite-dimensional equations of the form (50.1) in which the matrix Γ_0 can be diagonalized, only the Dirac equation has positive definite charge density and total charge.

2. Wave Equations With Positive Definite Energy Density.

Let us now try to find relativistically covariant wave equations of the form (50.1) which lead to positive definite energy densities.

The energy density is given by

$$w = \frac{i}{2} \left(\Gamma_0 \frac{\partial \psi}{\partial t}, \psi \right) - \left(\psi, \Gamma_0 \frac{\partial \psi}{\partial t} \right).$$

Replacing ψ in this equation by a plane wave in the rest system $\psi = \psi^{(s)} e^{-i p_0 t}$ and using (50.1), we obtain

$$w = \sum (\psi^{(0)}, \psi^{(0)}). \quad (52.7)$$

We, therefore, wish to find the conditions for which the quadratic form $(\psi^{(s)}, \psi^{(s)})$ is positive definite.

We shall assume that Γ_0 can be diagonalized. Then if $(\psi^{(s)}, \psi^{(s)})$ is positive,

$$(\Gamma_0 \psi, \Gamma_0 \psi) > 0, \quad (52.7')$$

where ψ is an arbitrary vector. To show this, we expand ψ in the eigenvectors $\psi_i^{(s)}$ on the left side of (52.7'),

$$\psi = \sum c_i \psi_i^{(s)},$$

obtaining

$$(\Gamma_0 \psi, \Gamma_0 \psi) = \sum |c_i|^2 \lambda_i^{(0)} (\psi_i^{(0)}, \psi_i^{(0)}) > 0$$

(the $\lambda_i^{(0)}$ are the eigenvalues of Γ_0).

We now treat (52.7') in a way analogous to that we used with (52.1). We shall merely present the results of this treatment here. It is found that (52.7') is positive definite only in two cases, namely when D_R contains the representations $\tau \sim (0, 0)$, $\tau \sim (1/2, 1/2)$ or the representations $\tau \sim (1, 0)$, $\tau \sim (0, 1)$, $\tau \sim (1/2, 1/2)$. The

possible values of l in the first and second cases are $l = 0$ and 1 . Therefore, the particles described by Equations (50.1) in this case have spin zero or spin one. It can be shown, however, that in the former case the matrix $\|c_{\tau\tau'}^k\|$ and in the latter case the matrix $\|c_{\tau\tau'}^s\|$ have no nonzero eigenvalues. According to the results of Section 51 it follows that in the first case the particles have spin zero, and in the second case spin one.

In the first case the representations D_R are five-dimensional, and in the second case ten-dimensional. It can be shown that in both cases the Γ_k matrices satisfy the conditions⁹⁾

$$\Gamma_i \Gamma_k \Gamma_i - \Gamma_i \Gamma_k \Gamma_i = -\delta_{ik} \Gamma_i - \delta_{ki} \Gamma_i. \quad (52.8)$$

If the spin is zero, the Γ_k are five-dimensional, whereas if the spin is one, the Γ_k are ten-dimensional.

Equations (50.1) for particles with spin zero can also be written in the form

$$\psi_k = \frac{\partial \psi}{\partial x_k}, \quad \frac{\partial \psi_k}{\partial x_k} = m^2 \psi. \quad (52.9)$$

Here it is clearly seen that the wave function consists of five components: the scalar ψ which transforms according to $\tau \sim (0, 0)$, and the vector $\frac{\partial \psi}{\partial x_k}$, which transforms according to $\tau \sim (1/2, 1/2)$.

If we eliminate the $\frac{\partial \psi}{\partial x_k}$ from (52.9), we obtain the well-known second-order scalar equation for ψ , namely,

$$\square \psi - m^2 \psi = 0. \quad (52.10)$$

Equations (50.1) for particles with spin one can be written

$$\psi_{ik} = \frac{\partial \psi_k}{\partial x_i} - \frac{\partial \psi_i}{\partial x_k}, \quad \frac{\partial \psi_{ik}}{\partial x_k} + m^2 \psi_i = 0. \quad (52.11)$$

It is clearly seen here that the wave function ψ contains ten components, namely the vector ψ_i , which transforms according to $\tau \sim (1/2, 1/2)$, and the antisymmetric tensor ψ_{ik} , which transforms according to $\tau \sim (1, 0)$ and $\bar{\tau} \sim (0, 1)$.

If the mass m is set equal to zero in (52.11), we obtain Maxwell's equations

⁹⁾ R. Duffin, Phys. Rev. 54, 1114 (1938).

$$F_{ik} = \frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}, \quad \frac{\partial F_{ik}}{\partial x_k} = 0. \quad (52.12)$$

We thus see that of all finite-dimensional equations of the form (50.1) in which Γ_k can be diagonalized, only the Dirac equation for spin $1/2$ particles leads to positive definite charge density and total charge, and only the equations for particles with spin 0 and 1 lead to positive definite energy density and total energy.

If Γ_k cannot be diagonalized, equations with positive definite charge or positive definite energy can be obtained for particles with spin greater than one.¹⁾

§ 53. Field Quantization. Spin and Statistics.

1. The Impossibility of Positive Definite Charge Density for Particles With Integral Spin.

We shall show that if a wave field describes particles with integral spin, then it is impossible to introduce a positive definite charge density or total charge, whereas if the field describes particles with half-integral spin, it is impossible to introduce a positive definite energy density or total energy.²⁾

In order to prove these assertions, let us divide quantities transforming according to irreducible representations $\tau \sim (j, k)$ of the Lorentz group into four classes: 1) the class $+1$ with integral j and k , 2) the class -1 with half-integral j and k , 3) the class $+e$ with integral j and half-integral k , and 4) the class $-e$ with half-integral j and integral k .

If some quantity ψ which transforms according to a reducible representation can be written as a sum of quantities belonging to the same class, then we shall say that ψ also belongs to this class.

It is easy to see that a tensor with an even (odd) number of indices belongs to class $+1$ (class -1). In particular, the wave vector k belongs to class -1 . It follows from (48.11) that a product of two quantities both of which belong to class $+1$ or class -1 can be separated only into quantities which belong to class $+1$; the product of a quantity in class $+1$ and a quantity in class -1 can be separated only into quantities of class -1 . Similarly, a product of two quantities each of which belongs to class $+e$ or $-e$, gives quantities only in class $+1$. The Table below gives the general rules for the classes of the products:

	$+1$	-1	$+e$	$-e$
$+1$	$+1$	-1	$+e$	$-e$
-1	-1	$+1$	$-e$	$+e$
$+e$	$+e$	$-e$	$+1$	-1
$-e$	$-e$	$+e$	-1	$+1$

These rules reduce to the relation $\epsilon^2 = +1$ and to the commutative law.

¹⁾ M. Fierz, Helv. Phys. Acta 12, 3 (1939); 13, 45 (1940).

²⁾ W. Pauli, Relativistic Theory of Elementary Particles, (Foreign Lit. Press, 1947) [the latter is a translation of an article by Pauli in Rev. Mod. Phys. 13, 203 (1941)].

We note further that if ψ transforms according to the representation $(\frac{1}{2}, k)$ which belongs to class +1 or class -1, then the complex conjugate ψ^* which transforms according to $(\frac{1}{2}, \frac{1}{2})$ belongs to the same class as ψ . If ψ belongs to class + ϵ (class - ϵ), then ψ^* belongs to class - ϵ (class + ϵ);

$$\begin{array}{ll} \psi^{++}, (\psi^{--})^* & \text{belongs to class} \quad +\epsilon, \\ \psi^{-+}, (\psi^{+-})^* & \text{belongs to class} \quad -\epsilon. \end{array} \quad (53.2)$$

Let us now consider a field corresponding to particles with integral spin. In this case $\frac{1}{2} + k$ is an integer, and therefore, ψ contains functions from class +1 and class -1. We shall assume that ψ is expanded in plane waves. Since the equations relating the components of ψ are invariant with respect to Lorentz transformation, they must be of the form

$$\sum k \psi^{(+)} = \sum \psi^{(-)}, \quad \sum k \psi^{(-)} = \sum \psi^{(+)}. \quad (53.3)$$

Indeed, only in this case will both parts of the equations, as can be seen from (53.1), transform in the same way under Lorentz transformation. Equations (53.3) will obviously remain invariant under the transformation

$$k_0 \rightarrow -k_0, \quad \psi^{(+)} \rightarrow \psi^{(-)}, \quad \psi^{(-)} \rightarrow -\psi^{(+)}. \quad (53.4)$$

Let us consider an arbitrary tensor T_{2n} of even order, which is quadratic with respect to ψ . Since such a tensor belongs to class +1, it can be written

$$T_{2n} \sim \sum \psi^{(+)} \psi^{(+)} + \sum \psi^{(-)} \psi^{(-)} + \sum \psi^{(+)} k \psi^{(-)}. \quad (53.5)$$

The transformation given by (53.4) leaves this form invariant. This is not true, however, for tensors T_{2n+1} of odd order (vectors, etc.). Such tensors belong to class -1 and therefore, are of the form

$$T_{2n+1} \sim \sum \psi^{(+)} k \psi^{(+)} + \sum \psi^{(-)} k \psi^{(-)} + \sum \psi^{(-)} \psi^{(+)}. \quad (53.6)$$

Under the transformation given by (53.4) these tensors change sign, so that $T_{2n+1} \rightarrow -T_{2n+1}$. In particular, this is true for the current vector j_μ . In other words, to every solution of the field equations there corresponds another solution for which the current has the opposite sign. Therefore, for particles of integral spin it is impossible to introduce a positive definite particle density which transforms like the fourth component of a vector.

We have thus shown the indefinite nature of the charge density, and therefore, of the total charge, in the case of integral-spin particles.

2. The Impossibility of Positive Definite Energy Density for Particles With Half-Integral Spin.

Let us now consider a field describing particles with half-integral spin. In this case $\frac{1}{2} + k$ is a half-integer and therefore, ψ belongs to class + ϵ and - ϵ . According to (53.1) and (53.2), the equations connecting the components of ψ in this case can be written

$$\begin{array}{l} \sum k \psi^{(+)} + \sum k (\psi^{(-)})^* = \sum \psi^{(-)} + \sum (\psi^{(+)})^*, \\ \sum k \psi^{(-)} + \sum k (\psi^{(+)})^* = \sum \psi^{(+)} + \sum (\psi^{(-)})^*. \end{array} \quad (53.7)$$

It is easy to see that these equations are invariant with respect to the transformation

$$\begin{array}{l} k_0 \rightarrow -k_0, \quad \psi^{(+)} \rightarrow i \psi^{(+)}, \quad (\psi^{(-)})^* \rightarrow i (\psi^{(-)})^*, \\ (\psi^{(+)})^* \rightarrow -i (\psi^{(+)})^*, \quad \psi^{(-)} \rightarrow -i \psi^{(-)}. \end{array} \quad (53.8)$$

Let us consider an arbitrary tensor T_{2n} of even order, which is quadratic in ψ . Since T_{2n} belongs to class +1, then according to (53.1) and (53.2) it can be written

$$\begin{aligned} T_{2n} \sim & \sum \psi^{(+)} \psi^{(+)} + \sum \psi^{(-)} \psi^{(-)} + \sum \psi^{(+)} k \psi^{(-)} + \\ & + \sum \psi^{(+)} (\psi^{(-)})^* + \sum \psi^{(-)} (\psi^{(+)})^* + \\ & + \sum (\psi^{(-)})^* k \psi^{(-)} + \sum (\psi^{(+)})^* k \psi^{(+)} + \sum (\psi^{(-)})^* k (\psi^{(+)})^* + \\ & + \sum (\psi^{(+)})^* (\psi^{(-)})^* + \sum (\psi^{(+)})^* (\psi^{(+)})^*. \end{aligned} \quad (53.9)$$

Similarly, an odd-order tensor T_{2n+1} which is quadratic in ψ belongs to class -1, and can be written

$$\begin{aligned} T_{2n+1} \sim & \sum \psi^{(+)} k \psi^{(+)} + \sum \psi^{(-)} k \psi^{(-)} + \sum \psi^{(-)} \psi^{(+)} + \sum \psi^{(+)} k (\psi^{(-)})^* + \\ & + \sum \psi^{(-)} k (\psi^{(+)})^* + \sum \psi^{(-)} (\psi^{(-)})^* + \sum \psi^{(+)} (\psi^{(+)})^* + \\ & + \sum (\psi^{(-)})^* k (\psi^{(-)})^* + \sum (\psi^{(+)})^* k (\psi^{(+)})^* + \sum (\psi^{(-)})^* (\psi^{(+)})^*. \end{aligned} \quad (53.10)$$

Applying (53.8) obviously does not change the form of T_{2n+1} , but changes the sign of T_{2n} , so that

$$T_{2n} \rightarrow -T_{2n}, \quad T_{2n+1} \rightarrow T_{2n+1}. \quad (53.11)$$

It follows from this that it is impossible to find a positive definite energy density, or therefore, a positive definite total energy, for particles with half-integral spin.

It does not follow from these premises that there always exists a positive definite energy density for particles with integral spin, or a positive definite charge density for particles with half-integral spin. On the contrary, it can be shown that only in the case of spin $1/2$ it is possible to introduce a positive definite charge density, and only in the case of spin 0 and 1 it is possible to introduce a positive definite energy density (it is assumed here that Γ_0 can be diagonalized).

3. Field Quantization for Integral and Half-Integral Spins.

From the above statements about positive definite particle density and energy density for the integral spin and half-integral spin cases one can derive an extremely important physical conclusion, according to which it is impossible to formulate a single-particle theory based on the Lorentz covariant equations obtained in Section 50. In other words, we should associate not a single particle, but a whole set of particles with the wave field we are considering. In order to do this it is necessary to consider the components of the wave function not only as functions of space and time, but also as certain operators acting in occupation-number space.

Since we are here studying noninteracting fields, we may start by considering single-particle states. Then, as is known, we must consider two possible cases, namely that in which there may be an arbitrary number of particles in each state, and that in which there can be no more than one particle in any state. We say that in the first case the particles satisfy Bose-Einstein statistics, and in the second case Fermi-Dirac statistics.¹⁾

In order that the field equations describe particles which satisfy one or the other type of statistics, the components of ψ must satisfy certain definite operator relations. These relations are called the quantum conditions, and the components of ψ which satisfy the quantum conditions are called the components of the quantized wave field. Generalizing the results of Sections 15 and 17, we assume that the commutators of the quantized wave field referring to bosons, and the anticommutators of the quantized wave field referring to fermions are certain c -numbers.²⁾

We shall henceforth give commutators and anticommutators the general name quantum brackets. Let us now prove the following fundamental theorem:³⁾ a wave field describing particles with integral spin should be quantized according to Bose-Einstein statistics; a wave field describing particles of half-integral spin should be quantized according to Fermi-Dirac statistics.

In proving this theorem, we shall first show the nature of the c -numbers given by the quantum brackets. These numbers may be coordinate-dependent scalars, vectors, tensors, etc., for various fields. We shall take as a postulate that all physical quantities commute if they refer to two events x' and x'' separated by a space-like interval, that is, for which $|x' - x''|^2 > |t' - t''|^2$. This postulate is in agreement with the principle of relativity, since in a time $|t' - t''|$ a perturbation which propagates from the first point with the velocity of light will travel a distance $|t' - t''|$ which is less than the distance $|x' - x''|$ between the points, and

¹⁾ Particles satisfying Bose-Einstein statistics are sometimes called bosons, and particles satisfying Fermi-Dirac statistics are sometimes called fermions.

²⁾ The commutator of the operators ψ_i and ψ_j is the expression

$$[\psi_i, \psi_j] = \psi_i \psi_j - \psi_j \psi_i,$$

and the anticommutator is the expression

$$\{\psi_i, \psi_j\} = \psi_i \psi_j + \psi_j \psi_i.$$

We shall denote the commutators and anticommutators by the brackets $[]_-$ and $[]_+$, respectively.

³⁾ This theorem is due to Pauli (see Footnote 2 on p. 521).

therefore, the second point is not effected by the perturbation from the first point. Therefore, measurements of two events which are separated by a space-like interval cannot be influenced by each other. It follows from this postulate that the c -numbers given by the quantum brackets must vanish if $|x' - x''|^2 - |t' - t''|^2 > 0$.⁴⁾

We shall now show how to find the desired c -numbers. Let us start by constructing a c -number scalar which vanishes for $|x' - x''|^2 - |t' - t''|^2 > 0$. Since the components of the quantized field ψ_k satisfy the second-order equation

$$\frac{\partial^2}{\partial x_k^2} \psi_k - m^2 \psi_k = 0, \quad (53.12)$$

as functions of space and time, the desired scalar must satisfy the same equation.

Since the scalar depends only on the combination $x^2 - t^2$, it clearly satisfies an ordinary second-order differential equation. It then follows that there exist only two linearly independent scalars depending on $x^2 - t^2$ and satisfying (53.12). The plane wave expansion of these scalars is of the form

$$\left. \begin{aligned} \Delta(x) &= \frac{1}{V} \sum_{\mathbf{p}} e^{i p x} \frac{\sin p_0 t}{p_0}, \\ \Delta^{(1)}(x) &= \frac{1}{V} \sum_{\mathbf{p}} e^{i p x} \frac{\cos p_0 t}{p_0}, \end{aligned} \right\} \quad (53.13)$$

where V is the normalizing volume, and $p_0 = \sqrt{\mathbf{p}^2 + m^2}$. Going to the limit $V \rightarrow \infty$ and replacing summation over $\mathbf{p}$ by integration over $V d^3 p / (2\pi)^3$, we obtain

$$\Delta(x) = \frac{1}{(2\pi)^3} \int d^3 p e^{i p x} \frac{\sin p_0 t}{p_0}, \quad \Delta^{(1)}(x) = \frac{1}{(2\pi)^3} \int d^3 p e^{i p x} \frac{\cos p_0 t}{p_0}. \quad (53.13')$$

The scalar character of these functions is seen from the fact that $d^3 p / p_0$ is an invariant; in fact,

$$\int \dots \frac{d^3 p}{p_0} = \frac{1}{2} \int \dots \delta(p^2 - p_0^2) d^4 p.$$

Since $\Delta(x) = 0$ when $t = 0$, the invariant function $\Delta(x)$ vanishes also when $x^2 - t^2 > 0$. In other words, $\Delta(x)$ vanishes for space-like intervals $x^2 - t^2 > 0$. On the other hand, the function $\Delta^{(1)}(x)$ does not satisfy this condition. Therefore, the desired scalar function can only be $\Delta(x)$.

The construction of c -number vectors and tensors which vanish for $x^2 - t^2 > 0$ presents no difficulty, since they can all be obtained by single or multiple differentiation of $\Delta(x)$ with respect to the coordinates x_α .

⁴⁾ We note that for $|x' - x''|^2 - |t' - t''|^2 > 0$, not only must the commutators vanish for fields corresponding to bosons, but so must the anticommutators for fields corresponding to fermions. If this were not true such quantities as the charge and current densities of fermions would not commute at space-like separated points.

Having gone through these preliminary remarks, let us consider the quantum bracket $[\psi_{\frac{1}{2}}(x'), \psi_{\frac{1}{2}}^*(x'')]_{\pm}$, where $\psi_{\frac{1}{2}}$ contains quantities from only one of the previously described classes $+1, -1, +\epsilon, -\epsilon$. According to the multiplication table (53.1) and relations (53.2), the product $\psi_{\frac{1}{2}}\psi_{\frac{1}{2}}^*$ belongs to class -1 if $\frac{1}{2} + \frac{1}{2}$ is a half-integer, or for half-integral spin, and it belongs to class $+1$ if $\frac{1}{2} + \frac{1}{2}$ is an integer, or for integral spin. Therefore, the quantum bracket must transform under Lorentz transformations like an odd-rank tensor in the case of half-integral spin, and like an even-rank tensor in the case of integral spin. In other words,

$$[\psi_{\frac{1}{2}}(x'), \psi_{\frac{1}{2}}^*(x'')]_{\pm} = Q_{2n+1} \Delta(x' - x'') \quad (53.14)$$

for half-integral spin, and

$$[\psi_{\frac{1}{2}}(x'), \psi_{\frac{1}{2}}^*(x'')]_{\pm} = Q_{2n} \Delta(x' - x'') \quad (53.15)$$

for integral spin, where Q_{2n+1} and Q_{2n} are differential operators containing, respectively, odd- and even-order derivatives with respect to x_{μ} , multiplied by certain constant coefficients. The explicit forms of Q_{2n+1} and Q_{2n} may be established if we know the wave equation of the field, the form of the energy-momentum tensor, and that of the current vector. The operators Q_{2n+1} and Q_{2n} are then found from the requirement that the energy and momentum of the field, as well as the charge, can be written in the form of sums over the energies, momenta, and charges of the individual particles (see Sections 15 and 17, in which the forms of Q_{2n+1} and Q_{2n} are established for the electromagnetic and electron-positron fields).¹⁾

Let us now consider the expression

$$X = [\psi_{\frac{1}{2}}(x'), \psi_{\frac{1}{2}}^*(x'')]_{\pm} + [\psi_{\frac{1}{2}}(x''), \psi_{\frac{1}{2}}^*(x')]_{\pm},$$

which is symmetric with respect to the events x' and x'' . Since $\Delta(x)$ is an even function of the space coordinates and an odd function of time (see (53.13)), it follows from the symmetry of X that X is a product of an even number of space derivatives of $\Delta(x' - x'')$, and an odd number of time derivatives. This is in agreement with Equation (53.14) for half-integral spin, but is in contradiction with (53.15) for integral spin, unless X vanishes.

We thus see that for integral spin

¹⁾ For particles with spin 0 and 1 the commutation rules (53.15) are of the form

$$i[\psi_{\frac{1}{2}}(x'), \gamma_{\lambda}(x'')] = \left[\left(\gamma_{\lambda} \frac{\partial}{\partial x_{\mu}} \right) \left(\gamma_{\mu} \frac{\partial}{\partial x_{\lambda}} - i m \right) \right]_{\lambda\mu} \Delta(x' - x''),$$

where $\chi = \psi^* \Delta$ and $\Delta = \|\alpha_{ik}\|$ is the matrix which defines the invariant bilinear form $(\varphi, \psi) = \epsilon_{ik} \alpha_{ik} \beta_{ik}$, $\varphi = \alpha_{ik} \xi_{ik}$, $\psi = \beta_{ik} \xi_{ik}$ (the ξ_{ik} are basis vectors). [V. Karman, J. Exptl.-Theoret. Phys. 21, 1337 (1951).]

$$[\psi_{\frac{1}{2}}(x'), \psi_{\frac{1}{2}}^*(x'')]_{\pm} + [\psi_{\frac{1}{2}}(x''), \psi_{\frac{1}{2}}^*(x')]_{\pm} = 0. \quad (53.16)$$

We have so far not considered the difference between Bose-Einstein and Fermi-Dirac statistics. If the particles are described by Bose-Einstein statistics, the commutators of the fields are c-numbers and the minus sign should be taken on the brackets in Equation (53.16); if, on the other hand, the particles are described by Fermi-Dirac statistics, then the anticommutators are c-numbers and the positive sign should be taken in Equation (53.16). If we take the positive sign on the brackets in (53.16), we arrive at an algebraic contradiction, since the left side of (53.16) is essentially positive for $x' = x''$. We have thus concluded that for integral spin, field quantization according to Fermi-Dirac statistics is impossible. In other words, particles with integral spin must always satisfy Bose-Einstein statistics.

As for the quantization of half-integral spin fields, such fields may be formally quantized according to Bose-Einstein statistics, but as we have seen above, the energy of the system will not be positive definite. It is clear from physical considerations that the energy of noninteracting particles must be considered positive. Therefore, half-integral spin fields must be quantized according to Fermi-Dirac statistics and we must require all (or almost all) states with negative energy to be filled. Then particles in positive energy states will not undergo transitions to occupied negative energy states.

If the state with all negative energies occupied is interpreted as the vacuum state of the half-integral spin field, then an unoccupied negative energy state can be considered as a state of a particle whose charge has the same absolute value as, but opposite sign from, a particle in a positive energy state (compare Section 17). We thus see that half-integral spin particles must satisfy Fermi-Dirac statistics.

II. BOUND STATE EQUATIONS.

§ 54. The Equation of Motion of an Electron in an External Field, With Radiative Corrections Taken Into Account.

1. The Method of Successive Approximations.

In Section 21, perturbation theory was applied to the equations of quantum electrodynamics in the Heisenberg representation, and equations for the S matrix were derived. We shall now show that by applying perturbation theory to the same equations we can obtain not only the elements of the S matrix, but also bound-state equations (the motion of a particle in a given field or the equation for interacting fields). The formal method used here is characterized by its great simplicity; it can also be used for constructing the S matrix.¹⁾

We shall start by considering the simple problem of electron motion in a given electromagnetic field.

Let us write the Dirac equation in the form

$$i\hat{\partial}\psi = ic\hat{\lambda}\psi, \quad (54.1)$$

where

¹⁾ A. Galanin, J. Exptl.-Theoret. Phys. 25, 448 (1952).

$$L = \hat{p} - i\hat{e}\hat{a} + m \quad (54.2)$$

$\hat{p}$ is the differential operator $\hat{p} = -i\gamma_\mu \frac{\partial}{\partial x_\mu}$, $\hat{a} = \gamma_\mu a_\mu$ (where a_μ is the given external electromagnetic potential), $\hat{A} = \gamma_\mu A_\mu$ (where A_μ is the electromagnetic radiation potential operator), and ψ is the electron-positron field operator.

We shall use the method of successive approximations, assuming $e\hat{A}\psi$ to be a small term in Equation (54.1). As for the external potential $\hat{a}$ the assumption that it is small is more conveniently used later.

The electromagnetic field operators A_μ satisfy the equation

$$\square A_\mu = -ie\bar{\psi}\gamma_\mu\psi. \quad (54.3)$$

Equation (54.1) can be written in the form of an equivalent integral equation

$$\psi = \psi_0 + L^{-1}(ie\hat{A})\psi, \quad (54.4)$$

where ψ_0 satisfies the homogeneous equation

$$L\psi_0 = 0. \quad (54.5)$$

The inverse operator L^{-1} is given by an infinite series in powers of $\hat{a}$ (it is assumed that this series has meaning), namely

$$L^{-1} = (\hat{p} + m)^{-1} + (\hat{p} + m)^{-1}(ie\hat{a})(\hat{p} + m)^{-1} + \dots \quad (54.6)$$

The inverse operator

$$(\hat{p} + m)^{-1} = (\hat{p} - m)(p^2 - m^2)^{-1} \quad (54.7)$$

is defined as follows. Let

$$f(x) = (2\pi)^{-4} \int e^{ikx} f_k d^4k,$$

and then

$$(\hat{p} + m)^{-1}f(x) = (2\pi)^{-4} \int \frac{i\hat{k} - m}{k^2 + m^2} e^{ikx} f_k d^4k, \quad (54.8)$$

where, in accordance with the explanation in Section 18, the integration over k_0 avoids the pole at $-k_0^2 + \mathbf{k}^2 + m^2 = 0$ by replacing m by $m - i\delta$ (where $\delta > 0$), and δ is then allowed to approach zero.

We shall solve the integral equation (54.4) by the method of successive approximations. In the first approximation, let us replace ψ_0 on the right side by ψ , so that

$$\psi = \psi_0 + L^{-1}(ie\hat{A})\psi_0. \quad (54.9)$$

Inserting (54.9) into the right side of (54.4), we obtain

$$\psi = \psi_0 + L^{-1}(ie\hat{A})\psi_0 + L^{-1}(ie\hat{A})L^{-1}(ie\hat{A})\psi_0. \quad (54.10)$$

Continuing this iteration procedure without limit, we obtain the infinite series

$$\psi = \psi_0 + L^{-1}(ie\hat{A})\psi_0 + L^{-1}(ie\hat{A})L^{-1}(ie\hat{A})\psi_0 + L^{-1}(ie\hat{A})L^{-1}(ie\hat{A})L^{-1}(ie\hat{A})\psi_0 + \dots$$

We may sum this series symbolically, writing

$$\psi = \psi_0 + L^{-1}(ie\hat{A})(L - ie\hat{A})^{-1}(ie\hat{A})\psi_0. \quad (54.11)$$

Applying the operator L to both sides, and bearing in mind (54.5), we obtain

$$L\psi = W\psi_0, \quad (54.12)$$

where

$$W = ie\hat{A} + (ie\hat{A})(L - ie\hat{A})^{-1}(ie\hat{A}). \quad (54.13)$$

2. Electromagnetic Vacuum Expectation Value.

In order to apply the equation obtained to concrete physical problems, we must find the expectation value of (54.12) in the electromagnetic vacuum state.

Before doing this, we must make several remarks as to the meaning of the operator ψ in Equation (54.1). The operator ψ is defined by the series

$$\psi = \sum a_n \psi_n(x); \quad \bar{\psi} = \sum a_n^\dagger \bar{\psi}_n(x),$$

where a_n is the annihilation operator for the electron in the state whose wave function is $\psi_n(x)$ and $a_n^\dagger$ is the electron-creation operator for the same state n . The functions ψ_n are a complete orthonormal set. If the calculation is performed in the interaction representation, it is convenient to choose the ψ_n as plane waves. Here, however, in using the Heisenberg representation, it is most convenient to choose the ψ_n as the set of functions that are the exact solutions of the problem under consideration. For instance, if we are dealing with the problem of the stationary states of the electron in the hydrogen atom, the ψ_n are the eigenfunctions for this problem, but with all radiative corrections taken into account (we may assume from physical considerations that such a set ψ_n exists). If, however, we wish to consider other problems such as the problem of electron motion in a given field accompanied by photon emission, then the ψ_n should be chosen as the exact solutions of this particular problem. In this case the index n would depend also on the state of the photon emitted by the electron. The action of the operator $a_n^\dagger$ on the occupation-number dependent wave function of the system would mean creation of an electron in state n in this case; but due to the equations of motion, the creation of such an electron is automatically associated with the creation of a photon. Thus, although the action of ψ on the wave function of the system involves only changing the occupation numbers for the electron, due to the equations of motion this action can be accompanied also by a change in the photon occupation numbers.

Let us now return to Equation (54.12), first treating the operator W only up to terms quadratic in $\hat{A}$:

$$L\psi = ie\hat{A}\psi_0 + (ie\hat{A})L^{-1}(ie\hat{A})\psi_0. \quad (54.14)$$

We shall consider the problem of stationary electron states in an external field (for instance, the hydrogen atom). We must, therefore, take the expectation value of (54.14) in the electromagnetic vacuum state. The left side will then not be altered. The expectation value of the first term on the right side vanishes, and the expectation value of the operator multiplying ψ_0 in the second term shall be denoted by

$$U_1 = \langle (ie\hat{A})L^{-1}(ie\hat{A}) \rangle. \quad (54.15)$$

We thus obtain the relation

$$L\psi = U_1\psi, \quad (54.16)$$

where we have replaced ψ_0 by ψ , on the right side, which is valid up to terms of order e^2 . We can now already treat ψ as a c-number. In calculating U_1 , we must bear in mind that the differential operator $\hat{p}$ contained in

L^{-1} acts on the function A_μ on the right of L^{-1} . If we move the coordinate-dependent part A_μ through L^{-1} from right to left, we must make use of the relations

$$\hat{p}e^{-ikx} = e^{-ikx}(\hat{p} - \hat{k}) \quad (54.17)$$

and

$$L^{-1}e^{-ikx} = e^{-ikx}(L - i\hat{k})^{-1}. \quad (54.18)$$

After the coordinate-dependent part of A_μ , which is represented by a Fourier integral, is moved through L^{-1} , we make use of the Fourier representation of the function D^- (see Section 16). We shall here assume that A_μ satisfies Equation (54.3) without the right side, i.e., that it satisfies the free-field equation, since the form of the function D^- was established just for a free field. We shall show below how the result changes when the right side of (54.3) is taken into account.

We will thus obtain the first part of the operator U_1 , which we shall denote by U_1^0 :

$$U_1^0 = -\frac{e^2}{16\pi^4} \int \gamma_\mu (\hat{p} - i\hat{k} - i\hat{a} + m)^{-1} \gamma_\mu k^{-2} d^4k, \quad (54.19)$$

where the integration over k_0 avoids the poles as has been indicated above.

Before calculating the second part of U_1 , which is due to the presence of the right side of Equation (54.3), let us examine Equation (54.19) in more detail. We note that in (54.19) the operator p_μ is a differential operator acting on $\hat{A}$, whereas in Section 42 p_μ was assumed to be a number.

Let us show that (54.19) gives the same result as the previous one in the first approximation in $\hat{A}$. To do this let us expand U_1 in a series, and keep only the first power of $\hat{A}$:

$$U_1^0 \approx U_{10}^0 + U_{11}^0, \quad (54.20)$$

where

$$U_{10}^0 = -\frac{e^2}{16\pi^4} \int \gamma_\mu (\hat{p} - i\hat{k} + m)^{-1} \gamma_\mu k^{-2} d^4k \quad (54.21)$$

and

$$U_{11}^0 = -\frac{e^2}{16\pi^4} \int \gamma_\mu (\hat{p} - i\hat{k} + m)^{-1} (i\hat{a}) (\hat{p} - i\hat{k} + m)^{-1} \gamma_\mu k^{-2} d^4k. \quad (54.22)$$

⁹ A. Abrikosov and I. Khalatnikov, J. Exptl.-Theoret. Phys., 21, 429 (1951).

Inserting U'_{11} into (54.16), we may consider ψ a plane wave, since accounting for the difference between ψ and a plane wave would mean considering the next power of the external field $\hat{a}$. Replacing $\hat{a}$ in (54.22) by a plane wave with wave vector $\hat{a}$, we obtain $(\bar{\psi}_2 | U'_{11} | \psi_1) = \bar{u}_2 e^{-i\hat{p} \cdot \hat{a}}; \psi_1 = u_1 e^{-i\hat{p} \cdot \hat{a}}; \hat{a} = \hat{a}_0 e^{-i\hat{p} \cdot \hat{a}}$

$$(\bar{\psi}_2 | U'_{11} | \psi_1) = \left(\bar{u}_2 | \frac{e^2}{16\pi^4} \int \gamma_p (\hat{p}_0 - i\hat{k} + \dots) + m^{-1} (ie\hat{a}_0) (\hat{p}_1 - i\hat{k} + m)^{-1} \gamma_p k^{-2} d^4k | u_1 \right), \quad (54.23)$$

where $\bar{u}_2$ and $\bar{u}_1$ are now numbers, and are equal to the electron momenta in states ψ_1 and ψ_2 . In this form, (54.23) agrees with the previously obtained result [see (42.1)].

If we insert U'_{11} into (54.16), we may not assume that ψ on the right side is a plane wave, but that $L\psi = 0$ for $\hat{a} \neq 0$. If we write ψ in a Fourier integral, we can consider the action of U'_{11} on one of the Fourier components ψ_p . Then the differential operator $\hat{p}$ can be replaced by a number (but $\hat{p}^2 + m^2 \neq 0$, since ψ_p is not a plane wave).

According to the method described in Sections 26 and 27, we must remove the infinities from U'_{11} . The renormalized operator U'_{11} (let us denote it by U'_{11R}) will have a second-order zero at $i\hat{p} = -m$. It follows from this that $U'_{11R}\psi_0$ is of the form

$$U'_{11R}\psi_0 = (i\hat{p} + m)f(\hat{p})(i\hat{p} + m)\psi_0 = (i\hat{p} + m)f(\hat{p})(i\hat{p} + m)\psi_0,$$

where $f(\hat{p})$ does not in general vanish for $i\hat{p} = -m$. When we calculate the matrix element of $U'_{11R}\psi_0$, we multiply on the left by $\bar{\psi}$, in this case this latter function must be considered a plane wave, since we are considering only linear terms in $\hat{a}$. We then obtain a vanishing result. Thus, the operator U'_{11} gives no contribution to (54.19), up to terms linear in $\hat{a}$.

If we want to calculate (54.19) with an accuracy up to $\hat{a}^2$, then we must add to (54.20) the following term of the power series expansion of (54.19) in $\hat{a}$:

$$U'_{12} = \frac{e^2}{16\pi^4} \int \gamma_p (\hat{p} - i\hat{k} + m)^{-1} (ie\hat{a}) (\hat{p} - i\hat{k} + m)^{-1} \gamma_p k^{-2} d^4k. \quad (54.24)$$

To this expression corresponds the diagram of Fig. 60.

In addition, when calculating $U'_{11}\psi_0$, we must write ψ_0 in the form

$$\psi_0 = \psi_0^{(0)} + (i\hat{p} + m)^{-1} (ie\hat{a}) \psi_0,$$

where $\psi_0^{(0)}$ is a plane wave. When calculating the matrix element $(\bar{\psi}_2 | U'_{11} | \psi_0)$, in addition to (54.23) we obtain the terms

$$(\bar{\psi}_2^{(0)} | U'_{11} (i\hat{p} + m)^{-1} (ie\hat{a}) | \psi_0^{(0)}) + (\bar{\psi}_2^{(0)} | (ie\hat{a}) (i\hat{p} + m)^{-1} U'_{11} | \psi_0^{(0)}), \quad (54.25)$$

(Fig. 61).

Finally, there remains a finite term also from U'_{10R} . We obtain

$$(\bar{\psi}_2 | U'_{10R} | \psi_0) = (\bar{\psi}_2 | (i\hat{p} + m)^{-1} f(\hat{p}) | \psi_0) = (\bar{\psi}_2^{(0)} | (ie\hat{a}) f(\hat{p}) (ie\hat{a}) | \psi_0^{(0)}). \quad (54.26)$$

Equation (54.26) corresponds to the diagram of Fig. 62. No simple renormalization diagrams such as that shown in Fig. 63 occur in this method of calculation.

Thus, by expanding (54.19) in a power series in $\hat{a}$, we obtain the same results as we obtained previously from the S-matrix theory. However, Equation (54.16) contains more, at least in principle. This lies in the fact that this equation gives the motion of an electron in the external field with the first-order radiative corrections taken into account, whereas from S-matrix theory we were able to obtain only the matrix elements of the effective potential U between states described by plane waves.

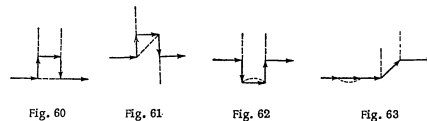


Fig. 60

Fig. 61

Fig. 62

Fig. 63

In the present problem there is, to a great extent, little difference between the two methods, but if we consider bound states of two interacting particles (rather than the motion of a single particle in a given field), then it is easier to use the present method to obtain the appropriate equations. Let us now return to the general equation (54.12). Let us take the expectation values of both sides of (54.12) in the electromagnetic vacuum state

$$| \psi \rangle = \langle W \rangle \psi_0. \quad (54.27)$$

In order to obtain an equation for ψ from (54.27), we must express ψ_0 in terms of ψ . This can be done by writing

$$\psi_0 = \psi - L^{-1} \langle W \rangle \psi_0$$

and applying the same iteration process as that which led to (54.12), considering ψ a known function, and ψ_0 unknown.

As a result we obtain

$$L\psi = U\psi, \quad (54.28)$$

where

$$U = \langle W \rangle - \langle W \rangle (L + \langle W \rangle)^{-1} \langle W \rangle. \quad (54.29)$$

If we write W in the form of a power series in $\hat{e}$, and use $\langle W \rangle$ only up to terms of some power of $\hat{e}$, then we see that the transition from $\langle W \rangle$ to U corresponds to eliminating from our considerations those many diagrams which can be divided into parts each of which is contained in a U operator of lower order in $\hat{e}$ and which are connected by only one electron line. This electron line is given by the operator $(\hat{p} - i\hat{e}\hat{a} + m)^{-1}$, which means that it may contain several vertices with the external field. We shall call such diagrams degenerate. They enter into the S matrix just as do other diagrams.

It is simple to see how degenerate diagrams appear in the S matrix.

Let us represent U in the form of a series $U = U_1 + U_2 + \dots$, where the index denotes the power of $\hat{e}$ in the operator. The function ψ can then be written in the series

$$\psi = \psi_0 + L^{-1}U_1\psi_0 + L^{-1}U_2\psi_0 + L^{-1}U_1L^{-1}U_1\psi_0 + \dots \quad (54.30)$$

When we calculate the matrix element $(\bar{\psi} | U | \psi)$ with the aid of (54.30), we obtain

$$(\bar{\psi} | U | \psi) = (\bar{\psi}_0 | U_1 | \psi_0) + (\bar{\psi}_0 | U_2 | \psi_0) + (\bar{\psi}_0 | U_1 L^{-1} U_1 | \psi_0) + \dots$$

The operators $U_1, L^{-1}U_1, U_2, L^{-1}U_2$, etc., are those that give the degenerate diagrams.

3. Vacuum Polarization. Electron-Positron Vacuum Expectation Value.

We shall now take account of the right side of Equation (54.3). In considering the single-electron problem (the two-electron problem will be considered in Section 55), we must find the expectation value of (54.3) in the electron-positron vacuum state. In the first approximation we may replace ψ in (54.3) by ψ_0 , so that $L\psi_0 = 0$, for $\hat{a} \neq 0$. Then the right side of (54.3) can be written in terms of the function $\hat{e}^2$. This function, differs, however, from that introduced in Section 18 in that it refers to an electron moving in the given external field $\hat{a}$. By making use of the relation between $\hat{e}^2$ and the inverse operator $(\hat{p} - i\hat{e}\hat{a} + m)^{-1}$, we can obtain the following expression for (54.3) in the first approximation in $\hat{e}^2$:

$$\square A_\mu = \frac{e^2}{16\pi^4} \int \text{Sp} [\gamma_\mu (\hat{p} - i\hat{k} - i\hat{e}\hat{a} + m)^{-1}] d^4k, \quad (54.31)$$

where the differential operator $\hat{p}$ acts only on $\hat{a}$.

Let us expand $(\hat{p} - i\hat{k} - i\hat{e}\hat{a} + m)^{-1}$ in a power series in $\hat{a}$. The term independent of $\hat{a}$ vanishes. In fact, if we write $(m - i\hat{k})^{-1}$ [we note that $\hat{p}$ must be considered zero if $\hat{a}$ is omitted from (54.31)] as a power

series in $\hat{a}$, the terms of odd power $\hat{a}$ vanish due to the symmetry in k -space and the terms of even power in k contain the trace of an odd number of γ matrices, so that they also vanish. The linear term in $\hat{a}$ gives the polarization current considered in Section 43. Indeed, if $\hat{a}$ is treated as a plane wave with propagation vector $\hat{a}$ (that is $\hat{a} = \hat{a}_0 e^{-i\hat{q}\hat{x}}$), then up to terms linear in $\hat{a}_0$ we obtain

$$\square A_\mu = \frac{e^2}{16\pi^4} \int \text{Sp} [\gamma_\mu (\hat{q} - i\hat{k} + m)^{-1} (i\hat{a}_0) (-i\hat{k} + m)^{-1}] d^4k e^{-i\hat{q}\hat{x}}. \quad (54.32)$$

Let us denote the polarization current by j_μ . The solution of the equation

$$\square A_\mu = -j_\mu$$

can be written in p -space in the form

$$A_\mu = A_\mu^0 - q^{-2} j_\mu, \quad (\square A_\mu^0 = 0). \quad (54.33)$$

We have inserted the first part of (54.33), that is A_μ^0 into (54.15) [making use of the fact that A_μ^0 satisfies the free-field equation and writing k^{-2} instead of D_μ^2 in (54.19)]. We must insert the second part of (54.33) into the first term on the right side of (54.14). As a result we obtain the second part of the operator U of (54.29), which is of first order in $\hat{e}^2$ and $\hat{a}$:

$$U_{11} = -\frac{e^2}{16\pi^4} \int \text{Sp} [\gamma_\mu (\hat{q} - i\hat{k} + m)^{-1} (i\hat{a}_0) (-i\hat{k} + m)^{-1}] d^4k, \quad (54.34)$$

which is in agreement with the result of Section 42 [see (42.2)].

In the further expansion of (54.31) in powers of $\hat{a}$, all terms of even order in $\hat{a}$ vanish by Purdy's theorem [if the current were written in the symmetric form $j_\mu = \frac{1}{2} (\bar{\psi} \gamma_\mu \psi - \bar{\psi}' \gamma_\mu \psi')$, where ψ' is the charge conjugate function, even powers of $\hat{a}$ would cancel automatically, which would prove Purdy's theorem; see also Section 24].

Successive approximations in powers of $\hat{e}^2$ are obtained by considering that ψ in (54.3) satisfies the exact equation (54.1). By generalizing (54.31) we obtain (for simplicity we set $\hat{a} = 0$)

$$\square A_\mu = \frac{e^2}{16\pi^4} \int \text{Sp} [\gamma_\mu (\hat{p} - i\hat{k} - i\hat{e}\hat{A} + m)^{-1}] d^4k. \quad (54.35)$$

The operator whose trace is being taken is defined as a power series in $\hat{A}$. If we take only the linear term in $\hat{A}$, then Equation (54.35) gives a linear equation for $\hat{A}$ (it is necessary, of course, to renormalize according to Sections 26 and 27).

In the momentum representation this equation is

$$p^2 \left[1 - \frac{e^2}{4\pi^2} F(p^2) \right] A_p = 0, \quad (54.36)$$

where according to (43.11)

$$F(p^2) = -\frac{4m^2 - 2p^2}{3p^2} \left(1 - \frac{0}{\tan \theta} \right) - \frac{1}{9}.$$

Since the equation for A_μ is now different [instead of $p^2 A_\mu = 0$, we have obtained (54.36)], the function D^E is also different. This means that in obtaining the expectation value, as for instance, in (54.15), we must replace k^{-2} by

$$k^{-2} \left[1 - \frac{e^2}{4\pi^2} F(k^2) \right]^{-1}$$

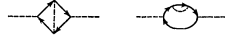


Fig. 64

Expanding $\left[1 - \frac{e^2}{4\pi^2} F(k^2) \right]^{-1}$ into a series in powers of e^2 , we obtain a sequence of terms corresponding to diagrams whose photon lines contain one, two, etc., successive simple closed loops.

The remaining terms of the power series in $\hat{A}$ give nonlinear expressions in (54.35). From these nonlinear terms one may again obtain linear equations for A_μ of the form (54.36), but then the function $F(p^2)$ will be more complicated. This is done by combining all but one of the factors $\hat{A}$ (according to Furry's theorem there is an odd number of them) into pairs, and replacing each pair by its electromagnetic vacuum expectation value. This gives rise to diagrams with more complicated closed loops (Fig. 64).

If this is not done, then (54.35) gives a nonlinear equation for A_μ , which describes processes such as scattering of light by light. When A_μ in (54.35) may be considered a c-number and a slowly varying function of the space-time coordinates, the operator on the right side of (54.35) can be calculated without an expansion in powers of $\hat{A}$. The result can be formulated as a change in the Lagrangian for the free electromagnetic field which is due to the interaction with the electron-positron vacuum, and is the same as that given in Section 47.

§ 55. The Equation of Motion of Two Interacting Electrons, with Radiative Corrections Taken Into Account.

1. The Equation of Motion of an Electron in a Real Photon Field.

We can now go on to the two-electron problem (we shall speak of two electrons, but the interaction may also be between two positrons or an electron and a positron). The interaction between electrons is accomplished by photon exchange. These photons are virtual, but they can be considered real photons from the point of view of a single electron. Therefore, we must first formulate an equation of motion, analogous to (54.28), in the field of an arbitrary number of real photons.

¹⁾ J. Schwinger, Phys. Rev. 82, 664 (1951).

Equation (54.1) for $a = 0$ and in the presence of real photons, can be written

$$L_r \psi = ie \hat{A}_r \psi, \quad (55.1)$$

where

$$L_r = \hat{p} + m - ie \hat{A}_r, \quad (55.2)$$

and $\hat{A}_r$ is a real photon field. Therefore, the desired equation is the same as (54.28) with the external field replaced by the real photon field. When taking the vacuum expectation value of some operator, for instance, one of the form of (54.15), we must interchange the photon emission operators in $\hat{A}$ and L_r . Clearly L_r does not commute with an emission (or absorption) operator for a photon whose momentum is the same as that of one of the real photons in $\hat{A}_r$. The contribution of these momenta in (54.15) approaches zero as the normalizing volume increases. We may, therefore, assume that L_r commutes with $\hat{A}$.

Let us now write out some of the terms of the expansion of U in powers of $\hat{A}_r$. For simplicity, we shall not write the brackets $\langle \dots \rangle$, but shall use $\hat{A}_r$ to denote the $\hat{A}$ operators for which the expectation value is taken. We shall also drop the index r on $\hat{A}_r$, so that $\hat{A}$ without a dot will denote a real photon field. Further, we shall include the term $ie \hat{A}_r$ of L_r in U . As a result we obtain

$$(\hat{p} + m) \psi = U \psi, \quad (55.3)$$

where

$$U = (ie \hat{A}) (\hat{p} + m)^{-1} (ie \hat{A}) + (ie \hat{A}) (\hat{p} + m)^{-1} (ie \hat{A}) (\hat{p} + m)^{-1} (ie \hat{A}) \times \\ \times (\hat{p} + m)^{-1} (ie \hat{A}) + \dots + ie \hat{A} + (ie \hat{A}) (\hat{p} + m)^{-1} \times \\ \times (ie \hat{A}) (\hat{p} + m)^{-1} (ie \hat{A}) + \dots + ie \hat{A} (\hat{p} + m)^{-1} (ie \hat{A}) + \dots \quad (55.4)$$

Here the first bracket contains no real photons, the second one contains one photon, the third one contains two, etc. The degenerate diagrams, for instance, the diagram corresponding to the term $(ie \hat{A}) (\hat{p} + m)^{-1} (ie \hat{A})$, are dropped in agreement with what has been said above.

2. The Equation of Motion of Two Interacting Electrons.

The wave function of a system of two electrons depends on the coordinates of two events and on two spinor indices which we shall suppress, this wave function can be written in the form

$$\psi(x_1, x_2) = \sum c_{ij} \psi_i(x_1) \psi_j(x_2), \quad (55.5)$$

where each function $\psi_i(x_1)$ or $\psi_j(x_2)$ satisfies Equation (55.3) with an operator U of the form of (55.4). Let us apply the operator $(\hat{p}_1 + m) \times (\hat{p}_2 + m)$ to both sides of (55.5), where $\hat{p}_1$ acts on x_1 and the first spinor index of $\psi(x_1, x_2)$, and $\hat{p}_2$ acts on x_2 and the second spinor index.

Since $\psi_1(x_1)$ and $\psi_2(x_2)$ satisfy Equation (55.3), we obtain³⁾

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(x_1, x_2) = U_1 U_2 \psi(x_1, x_2). \quad (55.6)$$

where U_1 is obtained from U of (55.4) by replacing $\hat{A}$ by $\hat{A}_1(x_1)$, and U_2 is obtained by replacing $\hat{A}$ by $\hat{A}_2(x_2)$.

Equation (55.6) can be solved by the method of successive approximations in the same way as we solved the equation for a single particle. We first obtain

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(x_1, x_2) = W_{12}\psi_0(x_1, x_2), \quad (55.7)$$

where

$$W_{12} = U_1 U_2 + U_1 U_2 [(i\hat{p}_1 + m)(i\hat{p}_2 + m) - U_1 U_2]^{-1} U_1 U_2, \quad (55.8)$$

and then the final equation

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(x_1, x_2) = U_{12}\psi(x_1, x_2), \quad (55.9)$$

where

$$U_{12} = \langle W_{12} \rangle - \langle W_{12} \rangle [(i\hat{p}_1 + m)(i\hat{p}_2 + m) + \langle W_{12} \rangle]^{-1} \langle W_{12} \rangle. \quad (55.10)$$

Here $\langle \rangle$ is the electromagnetic vacuum expectation value if there are no real photons in the problem, for instance, when considering the problem of stationary states of the system of two particles.

The transition from (55.8) to (55.10) involves the elimination of degenerate diagrams. These are now defined as those diagrams which can be divided into parts connected only by two electron lines belonging to different electrons. Examples of degenerate diagrams are shown in Fig. 65.

The final result for the operator U_{12} given by Equations (55.10), (55.8), and (55.4) can be written in a much more simple way, namely

$$U_{12} = \langle W_1 W_2 \rangle, \quad (55.11)$$

where W_1 and W_2 are obtained from (54.13) by replacing $\hat{A}$ by $\hat{A}_1(x_1)$ and $\hat{A}_2(x_2)$, respectively, and assuming that in taking the expectation value all degenerate diagrams are dropped.

Let us investigate Equation (55.9) in more detail for the simple case when

$$U_{12} = (ie\hat{A}_1(x_1))(ie\hat{A}_2(x_2)). \quad (55.12)$$

³⁾ Strictly speaking, Equation (55.6) describes the motion of particles which cannot annihilate. If the possibility of annihilation were taken into account, we would obtain a nonhomogeneous equation for $\psi(x_1, x_2)$.

We shall write the equation

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(x_1, x_2) = -e^2 \hat{A}_1(x_1) \hat{A}_2(x_2) \psi(x_1, x_2) \quad (55.13)$$

in the momentum representation. To do this, let us multiply both sides of (55.13) by $e^{-i(p_1 x_1 + p_2 x_2)}$, and integrate over x_1 and x_2 . The left side becomes

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(p_1, p_2),$$

where $\psi(p_1, p_2)$ is the Fourier transform of $\psi(x_1, x_2)$, and $\hat{p}_1$ and $\hat{p}_2$ are now matrices, rather than differential operators. The right side becomes

$$-e^2 \int \hat{A}_1(x_1) \hat{A}_2(x_2) e^{i(p_1 x_1 + p_2 x_2)} \psi(x_1, x_2) d^4 x_1 d^4 x_2.$$

Since

$$\hat{A}_1(x_1) \hat{A}_2(x_2) = -\frac{i}{16\pi^4} \int k^{-2} \gamma_{\mu}^{(1)} \gamma_{\nu}^{(2)} e^{-ik(x_1 - x_2)} d^4 k,$$

we finally obtain⁴⁾

$$(i\hat{p}_1 + m)(i\hat{p}_2 + m)\psi(p_1, p_2) = \frac{e^2}{16\pi^4} \int k^{-2} \gamma_{\mu}^{(1)} \gamma_{\nu}^{(2)} \psi(p_1 - k, p_2 + k) d^4 k, \quad (55.14)$$

where the indices (1) and (2) on γ denote action on the first and second spinor indices of ψ , respectively.

In principle these equations can be used to determine the energy levels of positronium. We shall not, however, go into this problem here.

The method of this section is easily generalized to more than two electrons.

III. MATHEMATICAL APPENDIX

§ 56. Calculation of Certain Integrals.

1. The Calculation of Integrals over a Finite Invariant Region.

We shall show how to calculate the integrals occurring in Sections 42 and 43. We shall perform the integration over a finite invariant region (N) which is characterized by the number N , and shall consider the limit

⁴⁾ E. Salpeter and H. Bethe, Phys. Rev. 84, 1232 (1951); A. Gelman, J. Exptl.-Theoret. Phys. 23, 448 (1952).

$N \rightarrow \infty$ to correspond to unbounded four-dimensional $\underline{k}$ -space. Such a finite invariant region may be defined, for instance, by the inequalities

$$|\underline{k}^2| \leq N^2, \quad \frac{|k_0|^2}{|\underline{k}^2|} \leq N^2, \quad (56.1)$$

where $\underline{k}$ is some time-like four-vector. In order to show that the region defined by (56.1) is finite, let us go to the reference system in which $\underline{k} = 0$. It then follows from the second inequality that $k_0^2 \leq N^2$, and from the first that $k^2 \leq 2N^2$. It is clear that the region of (56.1) will be finite in any other reference system, since the transition from one system to another is accomplished with the aid of a finite Lorentz transformation.

Let us now consider the integral

$$\int_{(N)} \frac{d^4 k}{(k^2 + i)^3}.$$

Since this integral converges over all $\underline{k}$ -space, we shall assume that $N \rightarrow \infty$. Let us first integrate over k_0 . In so doing we must bear in mind the rule of (18.34). Completing the contour of integration in the lower half-plane as shown in Fig. 66, we must find the residue of the integrand at the point $k_0 = +\sqrt{k^2 + i}$. This residue is $\frac{3}{16}(\underline{k}^2 + i) - \frac{1}{2}$, so that

$$\int_{-\infty}^{\infty} \frac{dk_0}{(k^2 - k_0^2 + i)^3} = \frac{3\pi i}{8} (k^2 + i)^{-3/2}.$$

Now integrating over the space components $\underline{k}$, we obtain

$$\int_{(N)} \frac{d^4 k}{(k^2 + i)^3} = \frac{\pi^2 i}{2} \frac{1}{N}, \quad N \rightarrow \infty.$$

Performing the substitutions $\underline{k} \rightarrow \underline{k} - \underline{p}$ and $l + p^2 \rightarrow l$, we arrive at

$$\int \frac{d^4 k}{(k^2 - 2pk + i)^3} = \frac{\pi^2 i}{2(l - p^2)}. \quad (56.2)$$

From symmetry considerations it is clear that

$$\int \frac{k_\alpha d^4 k}{(k^2 + i)^3} = 0.$$

Performing the substitutions $\underline{k} \rightarrow \underline{k} - \underline{p}$ and $l + p^2 \rightarrow l$ in this integral, we obtain

$$\int \frac{k_\alpha d^4 k}{(k^2 - 2pk + i)^3} = \frac{\pi^2 i p_\alpha}{2(l - p^2)}. \quad (56.3)$$

Let us now consider N finite but large compared with $|\underline{p}|$ and $|l|$, and let us integrate (56.2) and (56.3) over l :

$$\int_{(N)} \frac{d^4 k}{(k^2 - 2pk + i)^3} = -\pi^2 i \{ \ln(l - p^2) + A_0(N, p^2, s) \}, \quad (56.4)$$

$$\int_{(N)} \frac{k_\alpha d^4 k}{(k^2 - 2pk + i)^3} = -\pi^2 i p_\alpha \{ \ln(l - p^2) + B(N, p^2, s) \}. \quad (56.5)$$

Here A_0 is a logarithmically divergent constant which for large N behaves like $\ln N$ and B is in general a linearly divergent constant for $N \rightarrow \infty$, which, however, also diverges logarithmically for a symmetric region of integration. Differentiating (56.4) by p_α and comparing with (56.3), we obtain

$$\frac{\partial A_0}{\partial p_\alpha} = 0, \quad u = p^2,$$

so that $A_0(N, s)$ is a function only of N and s . Let us now differentiate (56.5) with respect to p_α :

$$\int_{(N)} \frac{k_\alpha k_\beta d^4 k}{(k^2 - 2pk + i)^3} = -\frac{\pi^2 i}{4} \delta_{\alpha\beta} \{ \ln(l - p^2) + B(N, u, s) \} + \frac{\pi^2 i}{2} p_\alpha p_\beta \left[\frac{1}{l - p^2} - \frac{\partial B}{\partial u} \right]. \quad (56.6)$$

Making use of the identity

$$\int_{(N)} \frac{k_\alpha k_\beta d^4 k}{(k^2 - 2pk + i)^3} = 2p_\alpha \int_{(N)} \frac{k_\beta d^4 k}{(k^2 - 2pk + i)^3} + l \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + i)^3} = \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + i)^3} \quad (56.7)$$

and inserting (56.6), (56.3), (56.2), and (56.4) into this expression, we obtain

$$\frac{u}{2} \frac{\partial B}{\partial u} + B = A_0 + \frac{1}{2},$$

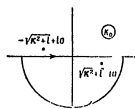


Fig. 66

whence

$$B = A_0(N, s) + \frac{1}{2} + \frac{1}{u^2} A'(N, s). \quad (58.8)$$

It follows from symmetry considerations that

$$\int_{(N)} \frac{k_s d^4 k}{(k^2 + l)^2} = 0. \quad (59.9)$$

Therefore, if we put $p_0 = 0$ in (56.5), we obtain

$$p_s B = 0,$$

whence it follows that $A'(N, s) = 0$ and

$$B(N, s) = A_0(N, s) + \frac{1}{2}. \quad (56.8')$$

Thus, Equations (56.4) - (56.6) can be written in the form

$$\left. \begin{aligned} \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + l)^2} &= -\pi^2 i \left\{ \ln(l - p^2) + A_0(N, s) \right\}, \\ \int_{(N)} \frac{k_s d^4 k}{(k^2 - 2pk + l)^2} &= -\pi^2 i p_s \left\{ \ln(l - p^2) + A_0(N, s) + \frac{1}{2} \right\}, \\ \int_{(N)} \frac{k_s k_t d^4 k}{(k^2 - 2pk + l)^2} &= \\ &= -\frac{\pi^2 i}{4} \delta_{st} \left\{ \ln(l - p^2) + A_0(N, s) + \frac{1}{2} \right\} + \frac{\pi^2 i}{2} \frac{p_s p_t}{l - p^2}. \end{aligned} \right\} \quad (56.10)$$

Now integrating the first of (56.10) over l , we obtain

$$\int_{(N)} \frac{d^4 k}{k^2 - 2pk + l} = \pi^2 i \left\{ (l - p^2) \ln(l - p^2) - (l - p^2) + A_0(N, s) l + C(N, s, u) \right\}. \quad (56.11)$$

Differentiating (56.11) with respect to p_0 ,

$$\int_{(N)} \frac{k_s d^4 k}{(k^2 - 2pk + l)^2} = -\pi^2 i p_s \left\{ \ln(l - p^2) - \frac{\partial C}{\partial u} \right\} \quad (56.11')$$

and comparing this expression with the second of Equations (56.10), we find that

$$\frac{\partial C}{\partial u} = -A_0 - \frac{1}{2},$$

whence

$$C = -u \left(A_0 + \frac{1}{2} \right) + A_2(N, s), \quad (56.12)$$

where A_2 diverges quadratically as $N \rightarrow \infty$. Finally, integrating the second of Equations (56.10) over l , we obtain

$$\int_{(N)} \frac{k_s k_t d^4 k}{k^2 - 2pk + l} = \pi^2 i p_s \left\{ (l - p^2) \ln(l - p^2) - (l - p^2) + \left(A_0 + \frac{1}{2} \right) l + \mathcal{E}(N, s, u) \right\}. \quad (56.13)$$

Differentiating this expression by p_t we obtain

$$\int_{(N)} \frac{k_s k_t d^4 k}{(k^2 - 2pk + l)^2} = \frac{\pi^2 i}{2} \delta_{st} \left\{ (l - p^2) \ln(l - p^2) - (l - p^2) + l \left(A_0 + \frac{1}{2} \right) + \mathcal{E}(N, s, u) \right\} - \pi^2 i p_s p_t \left[\ln(l - p^2) - \frac{\partial \mathcal{E}}{\partial u} \right]. \quad (56.14)$$

Making use of the identity

$$\int_{(N)} \frac{k_s k_t d^4 k}{(k^2 - 2pk + l)^2} - 2p_s \int_{(N)} \frac{k_s d^4 k}{(k^2 - 2pk + l)^2} + l \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + l)^2} = \int_{(N)} \frac{d^4 k}{k^2 - 2pk + l}$$

and Equations (56.10) - (56.12), and (56.14), we obtain

$$u \frac{\partial \mathcal{E}}{\partial u} + 2\mathcal{E} = -u \left(3A_0 + \frac{5}{2} \right) + A_2,$$

whence

$$\xi = -u \left(A_0 + \frac{5}{6} \right) + \frac{A_2}{2} + \frac{C'}{u^2}.$$

Putting $p_y = 0$ in (56.14), we find that $C' = 0$. Thus,

$$\left. \begin{aligned} C(N, s, u) &= - \left[A_0(N, s) + \frac{1}{2} \right] u + A_2(N, s), \\ \xi(N, s, u) &= - \left[A_0(N, s) + \frac{5}{6} \right] u + \frac{1}{2} A_2(N, s). \end{aligned} \right\} \quad (56.15)$$

We finally obtain the following expression for the integral of (56.14):

$$\int_{(N)} \frac{k_x k_y d^4 k}{(k^2 - 2pk + i)^2} = \frac{\pi^2 i}{2} \delta_{xy} \left\{ (l - p^2) \ln(l - p^2) - (l - p^2) + l \left(A_0 + \frac{1}{2} \right) - \right. \\ \left. - p^2 \left(A_0 + \frac{5}{6} \right) + \frac{1}{2} A_2 \right\} - \pi^2 i p_x p_y \left[\ln(l - p^2) + A_0 + \frac{5}{6} \right]. \quad (56.16)$$

2. Summary of the Integrals.

In conclusion we present a summary of the results.

$$\left. \begin{aligned} \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + i)^2} &= -\pi^2 i \left\{ \ln(l - p^2) + A_0(N) \right\}, \\ \int_{(N)} \frac{k_x d^4 k}{(k^2 - 2pk + i)^2} &= -\pi^2 i p_x \left\{ \ln(l - p^2) + A_0(N) + \frac{1}{2} \right\}, \\ \int_{(N)} \frac{k_x k_y d^4 k}{(k^2 - 2pk + i)^2} &= \frac{\pi^2 i}{2} \delta_{xy} \left\{ (l - p^2) \ln(l - p^2) - (l - p^2) + \right. \\ &\quad \left. + l \left[A_0(N) + \frac{1}{2} \right] - p^2 \left[A_0(N) + \frac{5}{6} \right] + \frac{1}{2} A_2(N) \right\} - \\ &\quad - \pi^2 i p_x p_y \left[\ln(l - p^2) + A_0(N) + \frac{5}{6} \right]. \end{aligned} \right\} \quad (56.17)$$

$$\left. \begin{aligned} \int_{(N)} \frac{d^4 k}{(k^2 - 2pk + i)^3} &= \frac{\pi^2 i}{2} \frac{1}{l - p^2}, \\ \int_{(N)} \frac{k_x d^4 k}{(k^2 - 2pk + i)^3} &= \frac{\pi^2 i}{2} \frac{p_x}{l - p^2}, \\ \int_{(N)} \frac{k_x k_y d^4 k}{(k^2 - 2pk + i)^3} &= -\frac{\pi^2 i}{4} \delta_{xy} \left[\ln(l - p^2) + A_0(N) + \frac{1}{2} \right] + \\ &\quad + \frac{\pi^2 i}{2} p_x p_y \frac{1}{l - p^2}. \end{aligned} \right\} \quad (56.18)$$

§ 57. L-Vectors and Spherical Functions.

1. Irreducible Tensors.

We have several times had to deal with quantities which transform under three-dimensional rotations according to irreducible representations of the rotation group. Such quantities are, for instance, eigenfunctions ψ_{LM} with eigenvalues L of the total angular momentum and with different eigenvalues M of the projection of this angular momentum, the components of the electric multipole moment Q_{LM} (which up to a constant factor are given by $\frac{L}{\sqrt{2L+1}} Y_{LM}$), etc. We shall briefly call such quantities L -vectors, where $2L+1$ is the number of components of this vector [and the transformation matrix is then $(2L+1)$ -dimensional]. In operating with L -vectors it is convenient to use a suitable algebra, which is to some degree analogous to tensor algebra.

We note that if L is an integer, the L -vector is equivalent to an ordinary three-dimensional tensor. In the general case a tensor of L -th rank has 3^L components, whose transformations are reducible. By the processes of symmetrization, antisymmetrization, and contraction one can separate out sets having less components, which transform under rotation independently of each other. Thus, for instance, of the nine components of a second-rank tensor T_{ik} we can consider its trace (a scalar), and its antisymmetric part, which consists of three quantities which transform like vector components (a first-rank tensor), after which we are left with the five independent components T_{ik}^0 of a symmetric tensor with trace zero:

$$T_{ik}^0 = \frac{1}{2} (T_{ik} + T_{ki}) - \frac{1}{3} T_{ii} \delta_{ik}.$$

The linear combinations of the T_{ik}^0 form an L -vector ($L=2$).

It is easily shown that the number of independent components of an L -th rank tensor in n -dimensional space $T_{ikl}^0 \dots$ (where $i, k, l, \dots = 1, 2, \dots, n$), symmetric with respect to all pairs of indices ($T_{ikl}^0 \dots = T_{k il}^0 \dots$, $T_{ikl}^0 \dots = T_{l ki}^0 \dots$, etc.) and vanishing for contraction of any pair ($T_{ii}^0 \dots = T_{kk}^0 \dots = \dots = 0$) is given by

$$N = \frac{(n+L-1)!}{(n-1)! L!} - \frac{(n+L-3)!}{(n-1)! (L-2)!}.$$

For three-space this equation is $N = 2L+1$; when $n=2$, we obtain $N=2$; when $n=4$, we obtain $N=(L+1)^2$.

We shall denote the components of an L -vector by $F_{\underline{L}}^{\underline{M}}$ (where $\underline{M} = -L, -L+1, \dots, L$). The basis in the $(2L+1)$ -dimensional space of the L -vector can be chosen so that the invariant bilinear form composed of the components of the L -vectors $\underline{F}$ and $\underline{G}$ (for the same $\underline{L}$), that is, their scalar product, be of the form

$$(FG) = \sum_{M=-L}^L (-1)^M F^M G^{-M}. \quad (57.1)$$

When $L = \frac{1}{2}$, the $F_{\underline{L}}^{\underline{M}}$ are the components of a spinor, and when $L=1$, they are the contravariant components of a vector in terms of the basis x_{μ} as defined in Section 3; that is,

$$F^0 = F_z; \quad F^{\pm 1} = \mp \frac{1}{\sqrt{2}} (F_x \mp iF_y)$$

We shall in general call F^M the contravariant components of an L -vector. If we define the covariant components by

$$F_M = (-1)^M F_{-M}, \quad (57.2)$$

then

$$FQ = \sum_M F^M G_M.$$

2. L -Vector Algebra.

The following problem is a fundamental one in L -vector algebra: we wish to use the components of the L -vector F^M and those of the L' -vector G^M to construct the L -vector H^M . This problem is solved by the representation theory of the rotation group; we shall here merely present the results. We note that if $L = L' = 1$, the problem reduces to constructing the vector product of two three-dimensional vectors which, as is well-known, is solved by

$$H_i = \epsilon_{ikl} F_k G_l, \quad (57.3)$$

where the indices i, k, l give the Cartesian components of the vectors, and ϵ_{ikl} is the unit antisymmetric tensor. Similarly, in the general case

$$H^M = C_{Mm}^{m'} F^m G^{m'}, \quad (57.4)$$

where $C_{Mm}^{m'}$ is a "third-rank tensor" whose components, like those of ϵ_{ikl} , do not depend on the choice of the coordinate system (the summation convention is used for m and m' from $-L$ to L and from $-L'$ to L' , respectively). In more detail, $C_{Mm}^{m'}$ is written with the indices $L, L',$ and L :

$$C_{Mm}^{m'} = C_{Lm}^{L'm'}.$$

We shall not present expressions for the $C_{Mm}^{m'}$ in general. For the case $L' = 1$ they are given in Section 4, and for $L' = \frac{1}{2}$ they are given in Section 10. We note that the $C_{Mm}^{m'}$ are nonzero only when

$$m' = M + m$$

$$|L - L'| \leq L \leq L + L'.$$

The latter condition is easily understood for integral L . From the products of tensors of the L -th and L' -th rank we can construct tensors of L -th rank, where the tensor of maximum rank is obtained by the direct (outer) product ($L = L + L'$), and that of minimum rank is obtained by contraction over all indices of the tensor of lower rank ($L = |L - L'|$).¹⁾

The $C_{Mm}^{m'}$ have the following orthogonality properties:

$$\left. \begin{aligned} \sum_{m, m'} C_{Lm}^{m'} C_{L'm'}^{m} &= \delta_{LL'} \delta_{mm'}, \\ \sum_{m, m'} C_{Lm}^{m'} C_{L'm'}^{m} &= \frac{2L+1}{2L'+1} \delta_{LL'} \delta_{mm'} \end{aligned} \right\} \quad (57.5)$$

and the following cyclic property²⁾

$$\frac{(-1)^{L-m}}{\sqrt{2L+1}} C_{Lm}^{m'} = \frac{(-1)^{L'-m'}}{\sqrt{2L'+1}} C_{L'm'}^{m} = (-1)^{L-L'} \frac{(-1)^{L-m}}{\sqrt{2L'+1}} C_{L'm'}^{m}, \quad (57.6)$$

We mention also the following important contraction relation:³⁾

$$\sum_{m, m'} C_{Lm}^{m'} C_{L'm'}^{m} = \omega C_{L'L}^{L'L}, \quad (57.7)$$

where ω depends only on L and L' , but not on M . [Equation (57.7) was used in Section 34 to sum over projections of the angular momenta, and to obtain an explicit expression for the correlation function (34.30).]

The vector multiplication rule (57.3), as is well-known, is sometimes written in a different form. One introduces the second-rank (antisymmetric) tensor G_{ik} which is the dual of G_k , given by

$$G_{ik} = \epsilon_{ikl} G_l,$$

and writes (57.3) in the form

$$H_i = F_k G_{ik}.$$

¹⁾ Equation (57.1) is a special case of (57.4) for $L = L'$, and $L = 0$.

²⁾ See L. Landau and E. Lifshitz, Quantum Mechanics, Section 97.

³⁾ G. Racah, Phys. Rev. 82, 439 (1942).

The multiplication rule (57.4) for $\underline{L}$ -vectors can be written similarly. We introduce the dual tensor or ($\underline{L}, \underline{L}$)-vector¹⁾

$$G_{Mm} = C_{Mm}^{M'}. \quad (57.8)$$

Then (57.4) becomes

$$H_M = F^{M'} G_{Mm}. \quad (57.9)$$

The tensor G_{Mm} is a rectangular matrix of $2\underline{L} + 1$ rows and $2\underline{L} + 1$ columns. We note that G_{Mm} depends not only on $\underline{L}$ and $\underline{L}$, but also on $\underline{L}'$.

From (57.4) or (57.9) we can, with the aid of the orthogonality properties (57.5), obtain an expression for the products of the components of the $\underline{L}$ -vector $\underline{F}$ and $\underline{G}$ in terms of the components of the $\underline{L}$ -vector $\underline{H}$ [this is the Clebsch-Gordon series, see (48.11)], namely

$$F_M G_m = \sum_{\underline{L}'} \rho_{\underline{L}'} H_{Mm} = \sum_{\underline{L}'} \rho_{\underline{L}'} \sum_{m'} C_{Mm'}^{L'm'} H_{m'}. \quad (57.10)$$

where the $\rho_{\underline{L}'}$ are independent of $\underline{M}, \underline{m}, \underline{m}'$.

3. Generalized Spherical Functions.

As we have already mentioned above, a three-dimensional vector is equivalent to an $\underline{L}$ -vector with $\underline{L} = 1$. The components of the unit vector $\underline{n}_{\underline{m}}$ (with $\underline{m} = 0, \pm 1$) are identical (up to a normalizing factor) with the spherical functions of first order $Y_{\underline{L}\underline{m}}$. Making use of the vector-multiplication law (57.4), we can use the components of $\underline{n}_{\underline{m}}$ to construct successively $\underline{L}$ -vectors with $\underline{L} = 2, 3$, etc. These will then be (up to a normalizing factor) the spherical functions $Y_{\underline{L}\underline{m}}$. We shall denote them, like the components of an $\underline{L}$ -vector, by $\underline{Y}_{\underline{M}}$ with the index $\underline{L}$ suppressed. Equation (57.4) then gives the recursion properties of spherical functions.

Equation (57.10) was used in Section 38 to express the tensor $\underline{n}_{\underline{m}}$ in terms of second-order spherical functions. Equations (57.4) and (57.10) can be used to obtain various relations for the spherical functions in a simple way. The coefficients $\rho_{\underline{L}'}$ are easy to find making use of special values such as $\underline{M} = 0$ and the angle $\theta = 0$.

Dual tensors [$(\underline{L}, \underline{L})$ -vectors] formed out of $\underline{L}$ -vectors by the spherical function $\underline{Y}_{\underline{Mm}}$ with $\underline{L} = 1, (\underline{L}' = \underline{L}, \underline{L} \pm 1)$ are identical with the spherical vectors $\underline{Y}_{\underline{L}\underline{M}}$ defined in Section 4. When $\underline{L} = \frac{1}{2}$ (and $\underline{L}' = \underline{L} + \frac{1}{2}$) the $\underline{Y}_{\underline{Mm}}$ are identical with the components of the spherical spinors $\underline{\Omega}_{\underline{L}\underline{M}}$ introduced in Section 10.

¹⁾ V. Berestetsky, A. Dolginov, and K. Ter-Martirosian, J. Exptl.-Theoret. Phys. 20, 527 (1950).

The wave function of a particle with spin $\underline{s}$ is an $\underline{s}$ -vector, and the wave functions of eigenstates of the angular momentum are the dual tensors of the spherical function $\underline{Y}_{\underline{Mm}}$ with $\underline{L} = \underline{s}$.¹⁾

The wave function for a particle with spin $\underline{s}$ in an eigenstate with angular momentum $\underline{L}$ can be written

$$\psi_{\underline{r}} = \sum_{M=-L}^L C^M Y_{M\underline{r}} \quad (\underline{r} = -s, \dots, s), \quad (57.11)$$

where the coefficients C^M form an $\underline{L}$ -vector (according to the meaning of an eigenstate with angular momentum $\underline{L}$, the wave function is determined by $2\underline{L} + 1$ independent coefficients).

¹⁾ We note that these wave functions can also be found directly by solving the eigenvalue problem for the infinitesimal rotation operator. In general these wave functions are expressed in terms of Jacobi polynomials; see I. Gelfand and Z. Shapiro, Prog. Math. Sci. 7, 3 (1952).